PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 13/00, C12N 1/21, 7/01
C12N 15/12, C12P 21/00

A1

(11) International Publication Number: WO 91/11461

(43) International Publication Date: 8 August 1991 (08.08.91)

(21) International Application Number: PCT/US91/00567

(22) International Filing Date: 28 January 1991 (28.01.91)

(30) Priority data:
470,888 26 January 1990 (26.01.90) US

(60) Parent Application or Grant
(63) Related by Continuation
US 470,888 (CIP)
Filed on 26 January 1990 (26.01.90)

(71) Applicant *(for all designated States except US)*: BIOGEN, INC. [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : PASEK, Mark, P. [US/US]; 151 G. Street, South Boston, MA 02127 (US). WINKLER, Gunther [AT/US]; 22 Grove Place, Unit 26, Winchester, MA 01890 (US). LIU, Theresa, R. [US/US]; 102 Emerson Road, Milton, MA 02186 (US).

(74) Agents: HALEY, James, F., Jr. et al.; Fish & Neave, 875 Third Avenue, 29th Floor, New York, NY 10022 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: C4 BINDING PROTEIN FUSION PROTEINS

(57) Abstract

This invention relates to multimeric and hetero-multimeric C4 binding protein (C4bp) fusion proteins and compositions and methods using them. More particularly, this invention relates to multimeric C4bp fusion proteins which are aggregates or assemblies of C4bp monomers linked to functional moieties. It also relates to C4bp fusion polypeptides and in particular CD4-C4bp fusion polypeptides comprising an amino acid sequence for a soluble human CD4 protein fused to a C4bp monomer having, preferably, four short consensus repeat regions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**ES**	Spain	**MG**	Madagascar
AU	Australia	**FI**	Finland	**ML**	Mali
BB	Barbados	**FR**	France	**MN**	Mongolia
BE	Belgium	**GA**	Gabon	**MR**	Mauritania
BF	Burkina Faso	**GB**	United Kingdom	**MW**	Malawi
BG	Bulgaria	**GN**	Guinea	**NL**	Netherlands
BJ	Benin	**GR**	Greece	**NO**	Norway
BR	Brazil	**HU**	Hungary	**PL**	Poland
CA	Canada	**IT**	Italy	**RO**	Romania
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		

C4 BINDING PROTEIN FUSION PROTEINS

TECHNICAL FIELD OF INVENTION

This invention relates to multimeric and hetero-multimeric C4 binding protein (C4bp) fusion proteins and compositions and methods using them. More particularly, this invention relates to multimeric C4bp fusion proteins which are aggregates or assemblies of C4bp monomers linked to functional moieties. It also relates to C4bp fusion polypeptides and in particular CD4-C4bp fusion polypeptides comprising an amino acid sequence for a soluble human CD4 protein fused to a C4bp monomer having, preferably, four short consensus repeat regions.

BACKGROUND OF THE INVENTION

In light of rapidly developing biotechnologies, researchers are producing novel delivery and carrier systems for pharmaceuticals, vaccines, diagnostics and other bioactive molecules. Optimally, these systems enhance the properties of the molecules they carry, complement those molecules with characteristics they lack and combine useful characteristics of different molecules. Of particular interest to researchers are the serum half-life of bioactive molecules, their affinity for target particles and cells, targetability of bioactive

molecules, bioactivity, immunogenicity and the ability to administer or deliver several molecules simultaneously.

Human C4 binding protein (hC4bp) is a molecule possessing many attractive characteristics as a delivery vehicle for bioactive molecules. Human C4bp is involved in the human complement system -- a group of immune system proteins whose functions include lysing invading cells, activating phagocytic cells and facilitating the clearance of foreign substances from the system. It regulates the action of proteins in this system, particularly C4 protein. Structurally, hC4bp is a flexible, disulfide-bonded molecule expected to have long serum half-life and the ability to target bioactive molecules to the lymph nodes. The serum form of hC4bp has a molecular weight of about 590 kD. On reducing SDS gels, hC4bp produces a strong band at about 70 kD, indicating a disulfide-bonded multimeric protein.

In the electron microscope, human C4bp appears as a structure with seven monomeric tentacles [B. Dahlback et al., "Visualization of Human C4b-Binding Protein and Its Complexes with Vitamin K-Dependent Protein S and Complement Protein C4b", Proc. Natl. Acad. Sci., USA, 80, pp. 3461-65 (1983)]. Although investigators have referred to the structure of human C4bp as spider-like, the flexibility of the tentacles of hC4bp renders that protein "octopus-like". Solution X-ray scattering experiments have suggested that in some environments, the tentacles of hC4bp may not be flayed out and the molecule may assume a compact shape [S.J. Perkins et al., "Unusual Ultrastructure of Complement-Component-C4b-Binding Protein of Human Complement by Synchroton X-Ray Scattering and

Hydrodynamic Analysis", Biochem. J., 233, pp. 799-807 (1986)].

A cDNA encoding the C4bp monomer has been cloned and characterized [L.P. Chung et al., "Molecular Cloning and Characterization of the cDNA Coding for C4b-Binding Protein of the Classical Pathway of the Human Complement System", Biochem. J., 230, pp. 133-41 (1985)]. Chung et al. refers to hC4bp as a polypeptide of 549 amino acids. The polypeptide predicted from the DNA sequence has a molecular weight of about 61.5 kD, rather than 70 kD as actually measured on reducing SDS gels. The difference in molecular weight apparently is due to glycosylation of the serum form of the polypeptide.

The first 491 amino acids from the N-terminus of the Chung et al. sequence are divisible into eight domains called short consensus repeat regions (SCRs) of about sixty amino acids each. These regions are designated, from N-terminus to C-terminus, SCR8 to SCR1. The SCR domains are defined by the amino acids of Figure 1 of this application as follows: SCR8 - +1 to +61; SCR7 - +62 to +123; SCR6 - +124 to +187; SCR5 - +188 to +247; SCR4 - +248 to +313; SCR3 - +314 to +374; SCR2 - +375 to +432; SCR1 - +433 to +491. These domains, which share significant sequence homology, each contain four similarly situated cysteine residues. These cysteine residues form intra-domain disulfide bonds in a regular pattern [J. Janatova et al., "Disulfide Bonds Are Localized Within the Short Consensus Repeat Units of Complement Regulatory Proteins: C4b-Binding Protein", Biochem., 28, pp. 4754-61 (1989)]. Within each SCR domain, the first cysteine residue bonds with the third and the second cysteine residue bonds with the fourth, forming a double-loop amino acid sequence. Thus, the SCRs are connected like

beads on a string. This pattern of intra-domain disulfide bonding is responsible for the conformational flexibility of the C4bp monomer.

In addition to the eight SCR domains, hC4bp also has a 58 amino acid sequence at the C-terminus, the C4bp core, which bears no homology to the other regions of the protein. This region is responsible for assembly of the molecule into a multimer. According to one model, the cysteine at position +498 of one C4bp monomer forms a disulfide bond with the cysteine at position +510 of another monomer.

In addition to seven C4bp monomers, human C4bp contains another subunit, a 45 kD polypeptide which is linked by disulfide bonds to the C4bp heptamer core [A. Hillarp and B. Dahlback, "Novel Subunit in C4b-Binding Protein Required for Protein S Binding", J. Biol. Chem., 263, pp. 12759-64 (1988)]. This subunit binds protein S, a protein involved in the regulation of blood clotting. When bound to protein S, protease C catalyses the transformation of clotting factors VIII and V from the active to inactive forms.

C4bp also exists in mammals other than humans. It has been isolated from both mouse and guinea pig [S.J. Lintin et al., "Derivation of the Sequence of the Signal Peptide in Human C4b-protein and Interspecies Cross-hybridization of the C4bp cDNA Sequence", FEBS Letters, 232, pp. 328-332 (1988)]. Analysis of mouse C4bp indicates that it contains contiguous SCRs, as does human C4bp. Mouse C4bp, however, has only six SCRs within each C4bp monomer and the multimer is held together by non-covalent bonds.

To date, the structural features of C4bp have not been utilized for the in vivo delivery of therapeutic or prophylactic agents. Despite advances in biotechnology, the need still exists for methods and

products which optimize the characteristics and delivery of pharmaceuticals, vaccines, diagnostics and bioactive molecules -- including polyvalency, affinity for a single target particle, serum half-life, bioactivity and, in some cases, immunogenicity.

SUMMARY OF THE INVENTION

The present invention solves these problems by providing multimeric and hetero-multimeric C4bp fusion proteins. Multimeric C4bp fusion proteins are aggregates or assemblies of C4bp monomers linked to functional moieties which may be pharmaceutical agents, vaccine agents, diagnostic agents or other bioactive molecules. Hetero-multimeric C4bp fusion proteins contain combinations of different C4bp monomers, different functional moieties, or combinations of both. This invention also provides multimeric and hetero-multimeric non-human C4bp fusion proteins.

C4bp fusion polypeptides comprise C4bp monomers fused or chemically coupled to a functional moiety. In particular, this invention provides the fusion polypeptide CD4(187)-C4bp(SCR4). This invention also relates to multimeric C4bp fusion proteins comprising monomeric C4bp fusion polypeptides. And this invention further provides DNA sequences encoding C4bp fusion polypeptides, recombinant DNA molecules comprising those sequences and unicellular host cells transformed with those molecules. This invention also provides methods for producing these fusion polypeptides by culturing such hosts. This invention also provides compositions comprising C4bp fusion polypeptides or proteins that are useful as therapeutic, prophylactic or diagnostic agents, particularly in diagnosing, preventing and treating AIDS, ARC and HIV infection.

The fusion proteins of this invention advantageously utilize various features of hC4bp, including its multimeric nature, its large size, the flexibility of its tentacles and its ability to migrate through the lymph nodes. Consequently, the bioactive molecules linked to C4bp monomers as functional moieties in such fusion proteins are characterized by one or more of the following: polyvalency, increased serum half-life, increased affinity for target particles or cells, greater bioactivity or immunogenicity and targetability.

Depending upon the choice of functional moiety, multimeric and hetero-multimeric C4bp fusion proteins according to this invention have many uses. Recognition molecules, such as those containing the antigen binding site of antibodies, viral receptors or cell receptors, are useful as functional moieties to target C4bp fusion proteins to particular antigens. When targeted in this manner, multimeric C4bp fusion proteins are useful to block the binding of viruses to cells, thereby preventing viral infection. C4bp fusion proteins may also be used to inhibit cell to cell binding such as that which characterizes pathologic inflammation. Due to the multivalency and conformational flexibility of the fusion proteins of this invention, we believe that they possess greater affinity for the target than monovalent or rigid multivalent molecules. In one embodiment of this invention, the functional moiety is the receptor on human lymphocytes, CD4, which is the target of the HIV virus -- the causative agent of AIDS and ARC.

When recognition molecules are used in conjunction with toxins, anti-retroviral agents or radionuclides in hetero-multimeric C4bp fusion proteins according to this invention, those proteins become

therapeutic agents which search out and destroy their target. C4bp fusion proteins having recognition molecules are also useful for signal enhancement in diagnostic assays. As large multimeric molecules, they present many binding sites for reporter molecules, such as horseradish peroxidase-conjugated antibodies. Alternatively, they may take the form of hetero-multimers, possessing both recognition molecules for the target and multiple reporter groups.

When the functional moiety component of the C4bp fusion protein is one or more immunogen from infectious agents, the proteins of this invention are useful in vaccines. And when the functional group is an enzyme, substrate, or inhibitor, the multimeric C4bp fusion proteins may function as agents with increased bioactivity.

The present invention also provides recombinant human C4bp and processes for production of that protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1C depict the DNA sequence and deduced amino acid sequence of human C4bp polypeptide derived from pJOD.C4bp.3. The negatively numbered amino acids correspond to the signal sequence, which is absent from the mature polypeptide. Throughout this application, references to C4bp by amino acid formula correspond to the coordinate system set forth in this figure.

Figure 2 depicts the structure of an SCR domain. It portrays an amino acid sequence of a short consensus repeat (SCR) region connected to adjacent SCRs. Each amino acid is represented by a circle. As described, infra, each SCR is held together by two disulfide bonds between cysteines 1 and 3 and between

cysteines 2 and 4, as depicted in this figure. The loop is depicted as the amino acid sequence between cysteines 1 and 4, inclusive. The joints are depicted as the amino acid sequences between two connected loops.

Figures 3A-3B depict the nucleotide sequence and deduced amino acid sequence of human CD4 protein. Nucleotides 1 to 636 are derived from pJOD.sCD4.Y187.SnaB1. Nucleotides 637 to 1377 are derived from p170.2. In this figure, the amino acids are numbered from -25 to 375. Throughout this application, references to CD4 by amino acid formula correspond to the coordinate system of this figure, unless otherwise specified.

Figure 4 depicts the domain structure of human CD4 protein. The numbered amino acids are cysteine residues involved in disulfide bonding according to Figures 3A-3B.

Figures 5A-5B depict the DNA sequences of oligomers C4bp.1 to C4bp.20, SCR.1, SCR.4, SCR.8, 312.20, 312.21, 312.35 and 312.36. In all sequences, left to right designates 5' to 3'.

Figures 6A-6H depict the construction of plasmids pJOD.C4bp and pJOD.sCD4.Y187.SnaB1.

Figure 7 depicts the construction of a plasmid containing a sequence encoding a CD4-C4bp fusion polypeptide according to this invention. A "CD4-C4bp fusion polypeptide" comprises amino acid sequences of human CD4 protein and C4bp. The top strand depicts pJOD.sCD4 including the adenovirus major late promoter (Ad MLP); the 5' untranslated sequence (5' UTS); the ATG initiation codon and signal sequence encoding region; the region encoding human CD4 protein through the codon for tyrosine (TAC(187)); the SnaBI site (TACGTA); the BglII site (AGATCT); and the SV40

polyadenylation control sequence. The bottom strand depicts pJOD.C4bp, including the region encoding SCR8-SCR1, the core and termination codon and the MIS gene polyadenylation control sequence.

Figure 8 depicts the results of purification of recombinant human C4bp (rhC4bp) and the serum form of human C4bp (serum) by HPLC.

Figure 9 depicts illustrative embodiments of C4bp fusion polypeptides and proteins according to this invention.

Figure 10 is a table summarizing the antibodies used in ELISA assays 1-9, described herein.

DETAILED DESCRIPTION OF THE INVENTION

In order that the invention herein described may be more fully understood, the following detailed description is set forth.

In the description, the following terms are employed:

"C4 binding protein" ("C4bp") refers to a polypeptide having the amino acid sequence depicted in Figure 1 from amino acids -32 to +549. It should be understood that expression of polypeptides often involves post-translational modifications, such as cleavage of the signal sequence, intramolecular disulfide bonding, glycosylation and the like. Accordingly, the term, C4 binding protein, also contemplates such modifications to the amino acid sequence of C4bp. It also encompasses naturally occurring genetic polymorphisms. The term also includes C4 binding proteins from natural, recombinant or synthetic sources.

"Multimeric C4bp fusion proteins" and "hetero-multimeric C4bp fusion proteins" each comprise aggregates or assemblies of C4bp fusion polypeptides.

"C4bp fusion polypeptides" comprise a C4bp monomer bound to a functional moiety. "Functional moieties" may be polypeptides ("polypeptide moieties") or chemical compounds ("chemical moieties"). One may produce multimeric C4bp fusion proteins by genetic fusion, chemical synthesis, or chemical coupling techniques.

When the functional moiety is a polypeptide, genetic fusion is preferred. This involves, for example, creating a hybrid DNA sequence encoding the C4bp fusion polypeptide in which the 3' end of a DNA sequence encoding the polypeptide is ligated to the 5' end of a DNA sequence encoding a C4bp monomer. Upon expression in an appropriate host, this hybrid DNA sequence will produce a C4bp fusion polypeptide that will assemble into a multimer.

A "C4bp monomer" as used herein is a polypeptide comprising a C4bp core or, more preferably, a sequence of at least one SCR fused to the N-terminus of a C4bp core. A "C4bp core" encompasses, at a minimum, amino acids +498 to +549 of Figure 1 and, preferably, amino acids +492 to +549. As used herein, an "SCR" is a polypeptide fragment of C4bp. An SCR comprises, at a minimum, a loop and, at a maximum, a loop and two joints. A "loop" comprises the amino acid sequence encompassed by the first and fourth cysteines of the eight SCR domains as defined above. That is, the loops encompass amino acids +2 to +60 of SCR8, +65 to +122 of SCR7, +127 to +186 of SCR6, +191 to +246 of SCR5, +251 to +312 of SCR4, +316 to +375 of SCR3, +378 to +432 of SCR2 and +446 to +490 of SCR1. A "joint" comprises the amino acid sequences between and (in the cases of SCR8 and SCR1) outside the loops. Thus, each loop is bonded to another loop via a joint. SCRs having joints are preferable to those that do not have

joints because it is unlikely that two loops bonded without a joint will be as flexible as those bonded through a joint. Most preferably, an SCR comprises the amino acid sequence of the SCR domains defined above. It should be understood that one could make minor alterations in the amino acid sequence of an SCR, for example by adding a few amino acids to the short loops of SCR1 and SCR8.

The C4bp monomers of this invention include any sequence of SCRs, including SCRs strung together at random. However, it is an object of this invention to produce proteins least likely to evoke an immune response against the C4bp monomer. Therefore, more preferably, the amino acid sequence of the C4bp monomer corresponds to at least a fragment of the amino acid sequence of mature C4bp, which is not normally immunogenic. Thus, the C4bp monomer, C4bp(SCR8) corresponds to the mature C4bp polypeptide. C4bp(SCR4) corresponds to amino acids +248 to +549 of Figure 1. C4bp(SCR1) corresponds to amino acids +433 to +549 of Figure 1. The C4bp monomer, C4bp(SCR4), is most preferable.

According to alternate embodiments of this invention, C4bp monomers include variable numbers of SCRs. At a minimum, there may be no SCRs. At a maximum, C4bp monomers may contain about 32 SCRs, about as many as the longest known repeating unit molecule, CR1, which has 30 domains [L.B. Klickstein, "Isolation of NH2-terminal CR1 (CD35) Clones and Expression of Recombinant Human CR1", FASEB J., 2, p. A1832 #8921 (1988)].

A C4bp monomer containing more than eight SCRs corresponds more preferably to the amino acid sequence of mature C4bp fused to at least a fragment of

the same. For example, a sixteen SCR monomer may comprise SCR8-SCR1 fused to SCR8-SCR1.

DNA sequences encoding C4bp monomers are derived from DNA sequences encoding C4bp. Several methods are available to obtain these DNA sequences. First, one can chemically synthesize the C4bp gene or a degenerate version of it using a commercially available chemical synthesizer. We have presented a DNA sequence for C4bp in Figure 1, including the signal sequence from nucleotides +1 to +96. It confirms the sequences presented by Chung et al., supra, and Lintin et al., supra, except for three silent nucleotide substitutions. The differences are at the codons beginning at nucleotides 625, 1402 and 1459, which read GGC, TGG and GAG, respectively. Chung et al. identifies those codons as GGT, TGC and GAA.

Second, one can isolate a cDNA sequence encoding the C4bp polypeptide by screening a cDNA library. Many screening methods are known to those of skill in the art. For example, one can screen colonies by nucleic acid hybridization with oligonucleotide probes. Probes can be prepared by chemically synthesizing an oligonucleotide having part of the known DNA sequence of C4bp. Alternatively, one can construct cDNA libraries in expression vectors, such as λgt11, and screen the colonies with anti-hC4bp antibodies.

Third, one can isolate a cDNA encoding C4bp by amplifying mRNA using the polymerase chain reaction (PCR). We describe this process in Example I.

The DNA sequence encoding the C4bp monomer may then be fused to a DNA sequence encoding a functional moiety, such as a polypeptide moiety. DNA sequences for polypeptides useful in this invention are available from many sources. These include DNA

sequences described in the literature and DNA sequences encoding particular polypeptides obtained by any of conventional molecular cloning techniques.

This invention also contemplates non-human C4bp fusion proteins comprising non-human C4bp fusion polypeptides. In such fusion polypeptides, the C4bp monomers comprise C4bp cores and SCRs derived from the amino acid sequence of a non-human C4bp. Any non-human C4bp having monomeric units that assemble into a multimer are useful for this purpose. Such C4bp multimers exist in the guinea pig and mouse [Lintin et al., supra]. Mouse C4bp is preferable, because its amino acid sequence is known to contain contiguous SCRs.

A wide array of polypeptides are useful to produce the C4bp fusion proteins or fusion polypeptides of this invention. Those most useful include polypeptides that are advantageously administered in multimeric form. For example, viral receptors or cell receptors or ligands are useful, because they typically bind to particles or cells exhibiting many copies of the receptor. Fusion proteins containing these polypeptides are useful in therapies that involve inhibiting viral-cell or cell-cell binding. Useful viral-cell receptors include ICAM1, a rhinovirus receptor; the polio virus receptor [J. White and D. Littman, "Viral Receptors of the Immunoglobulin Superfamily", Cell, 56, pp. 725-28 (1989)] and, most preferably, CD4, the HIV receptor. Cell-cell receptors or ligands include members of the vascular cell adhesion molecule family, such as ICAM1, ELAM1, and VCAM1 and VCAM1b and their lymphocyte counterparts (ligands); the lymphocyte associated antigens, LFA1, LFA2 (CD2) and LFA3, members of the CD11/CD18 family, and VLA4. These molecules are involved in pathologic

inflammation [M.P. Bevilacqua et al., "Identification of an Inducible Endothelial-leukocyte Adhesion Molecule", Proc. Natl. Acad. Sci., USA, 84, pp. 9238-42 (1987); L. Osborn et al., "Direct Expression Cloning of Vascular Cell Adhesion Molecule 1: A Cytokine-induced Endothelial Protein that Binds to Lymphocytes," Cell, 59, pp. 1203-11 (1989) and Hession et al., WO 90/13300].

Bacterial immunogens, parasitic immunogens and viral immunogens are useful as polypeptide moieties to create multimeric or hetero-multimeric C4bp fusion proteins useful as vaccines. Bacterial sources of these immunogens include those responsible for bacterial pneumonia and pneumocystis pneumonia. Parasitic sources include malarial parasites, such as Plasmodium. Viral sources include poxviruses, e.g., cowpox virus and orf virus; herpes viruses, e.g., herpes simplex virus type 1 and 2, B-virus, varicella-zoster virus, cytomegalovirus, and Epstein-Barr virus; adenoviruses, e.g., mastadenovirus; papovaviruses, e.g., papillomaviruses, and polyomaviruses such as BK and JC virus; parvoviruses, e.g., adeno-associated virus; reoviruses, e.g., reoviruses 1, 2 and 3; orbiviruses, e.g., Colorado tick fever; rotaviruses, e.g., human rotaviruses; alphaviruses, e.g., Eastern encephalitis virus and Venezuelan encephalitis virus; rubiviruses, e.g., rubella; flaviviruses, e.g., yellow fever virus, Dengue fever viruses, Japanese encephalitis virus, Tick-borne encephalitis virus and hepatitis C virus; coronaviruses, e.g., human coronaviruses; paramyxoviruses, e.g., parainfluenza 1, 2, 3 and 4 and mumps; morbilliviruses, e.g., measles virus; pneumovirus, e.g., respiratory syncytial virus; vesiculoviruses, e.g., vesicular stomatitis virus; lyssaviruses, e.g., rabies virus; orthomyxoviruses,

e.g., influenza A and B; bunyaviruses e.g., LaCrosse virus; phleborviruses, e.g., Rift Valley fever virus; nairoviruses, e.g., Congo hemorrhagic fever virus; hepadnaviridae, e.g., hepatitis B; arenaviruses, e.g., lcm virus, Lassa virus and Junin virus; retroviruses, e.g., HTLV I, HTLV II, HIV I and HIV II; enteroviruses, e.g., polio virus 1, 2 and 3, coxsackie viruses, echoviruses, human enteroviruses, hepatitis A virus, hepatitis E virus, and Norwalk virus; rhinoviruses e.g., human rhinovirus; and filoviridae, e.g., Marburg (disease) virus and Ebola virus.

More specifically, this invention provides C4bp fusion polypeptides comprising a polypeptide moiety comprising viral polypeptides having hepatitis B virus e antigenicity. A DNA sequence encoding hepatitis B virus e antigens ("HBeAg") is described in L. Mimms et al., "Production, Purification, and Immunological Characterization of a Recombinant DNA-derived Hepatitis B e Antigen", Viral Hepatitis and Liver Disease, pp. 248-251 Alan R. Liss, Inc. (1988). The amino acids encoded by this sequence correspond to the amino-terminal 144 amino acids of Hepatitis B Virus core antigen ("HBcAg") (subtype adw_2). Alternatively, a DNA sequence encoding HBeAg includes the sequence corresponding to amino acids 1 to 144 of HBcAg, as set forth in M. Pasek et al., "Hepatitis B Virus Genes and Their Expression in E. coli", Nature, 282, pp. 575-579 (1979). A DNA sequence which encodes HBeAg may also be obtained according to the processes set forth in Murray et al., U.S. patent 4,758,507. We shall refer herein to a DNA sequence encoding, or a polypeptide having, HBeAg amino acids numbers 2 (Asp) to X as "HBeAg(2-X)".

An immunoglobulin or fragment thereof that binds to a target molecule is also useful as a

functional moiety. Immunoglobulin molecules are bivalent, but immunoglobulin-C4bp fusion proteins will be multivalent and may demonstrate increased affinity or avidity for the target. It has been demonstrated that single domain antibodies (dAbs) are useful [E.S. Ward et al., "Binding Activities of a Repertoire of Single Immunoglobulin Variable Domains Secreted from Escherichia coli," Nature, 341, pp. 544-46 (1989)]. One can generate monoclonal Fab fragments recognizing specific antigens using the technique of W.D. Huse et al. and use individual domains as functional moieties in multimeric or hetero-multimeric C4bp fusion proteins according to this invention [W.D. Huse et al., "Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda," Science, 246, pp. 1275-81 (1989)]. (See also A. Skerra and A. Pluckthun, "Assembly of a Functional Immunoglobulin Fv Fragment in Escherichia coli", Science, 240, pp. 1038-43 (1988)).

One may also produce multimeric C4bp fusion proteins as agents with increased bioactivity when the functional moiety is an enzyme, enzyme substrate or enzyme inhibitor. We expect such agents to exhibit increased bioactivity because multimers have a higher density of the moiety and will exhibit increased turnover rate. For example, a multimeric C4bp fusion protein with tissue plasminogen activator would have greater clot-dissolving catalytic activity than its monovalent counterpart. Multimeric C4bp fusion proteins with hirudin, C-terminal hirudin peptides (described in PCT patent application WO 90/03391, incorporated herein by reference) and molecules based on hirudin structure (i.e., hirulogs, described in U.S. patent application 549,388, filed July 6, 1990, incorporated herein by reference) may display greater

anti-coagulant activity than monomers of these polypeptides.

Other useful functional moieties include polypeptides such as cytokines, including the various IFN-α's, particularly α2, α5, α8, IFN-β and IFN-γ, the various interleukins, including IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7 and IL-8 and the tumor necrosis factors, TNF-α, and β. In addition, functional moieties include, for example, monocyte colony stimulating factor (M-CSF), granulocyte colony stimulating factor (G-CSF), granulocyte macrophage colony stimulating factor (GM-CSF), erythropoietin, platelet-derived growth factor (PDGF) and human and animal hormones, including growth hormones and insulin.

According to one embodiment of this invention, multimeric C4bp fusion proteins comprise CD4-C4bp fusion polypeptides. CD4 is the receptor on those white blood cells, T-lymphocytes, which recognize HIV, the causative agent of AIDS and ARC [P.J. Maddon et al., "The T4 Gene Encodes the AIDS Virus Receptor and Is Expressed in the Immune System and the Brain", Cell, 47, pp. 333-48 (1986)]. Specifically, CD4 recognizes the HIV viral surface protein, gp120/160. In these fusion polypeptides, the functional moiety is a polypeptide moiety comprising CD4 or a fragment thereof, preferably soluble CD4.

The nucleotide sequence and a deduced amino acid sequence for a DNA that encodes the entire human CD4 protein have been reported [P.J. Maddon et al., "The Isolation and Nucleotide Sequence of a cDNA Encoding the T Cell Surface Protein T4: A New Member of the Immunoglobulin Gene Family", Cell, 42, pp. 93-104 (1985); D.R. Littman et al., "Corrected CD4 Sequence", Cell, 55, p. 541 (1988)]. Based upon its

deduced primary structure, the CD4 protein is divided into functional domains as follows:

Structure/Proposed Location	Amino Acid Coordinates In Figures 3A-3B
Hydrophobic/Secretory Signal	-25 to -1
First Immunoglobulin-related domain/Extracellular	+1 to +107
Second Immunoglobulin-related domain/Extracellular	+108 to +177
Third Immunoglobulin-related domain/Extracellular	+178 to +293
Fourth Immunoglobulin-related domain/Extracellular	+294 to +370
Hydrophobic/Transmembrane Sequence	+371 to +391
Very Hydrophilic/ Intracytoplasmic	+392 to +431

The first immunoglobulin-related domain can be further resolved into a variable-related (V) region and joint-related (J) region, beginning at about amino acid +95. [S.J. Clark et al., "Peptide and Nucleotide Sequences of Rat CD4 (W3/25) Antigen: Evidence for Derivation from a Structure with Four Immunoglobulin-related Domains", Proc. Natl. Acad. Sci., USA, 84, pp. 1649-53 (1987)].

These domains also correspond roughly to structural domains due to intra-domain disulfide bonding. Thus, disulfide bonds join amino acids at positions +16 and +84 in the first immunoglobulin-related domain, amino acids +130 and +159 of the second immunoglobulin-related domain and amino acids +303 and +345 of the fourth immunoglobulin-related domain. Figure 4 depicts the domain structure of the human CD4 protein of Figures 3A-3B.

Soluble CD4 proteins have been constructed by truncating the full length CD4 protein at amino acid +375, to eliminate the transmembrane and cytoplasmic domains. Such proteins have been produced by recombinant DNA techniques and are referred to as recombinant soluble CD4 (rsCD4) [R.A. Fisher et al., "HIV Infection Is Blocked In Vitro by Recombinant Soluble CD4", Nature, 331, pp. 76-78 (1988); Fisher et al., PCT patent application WO 89/01940 (incorporated herein by reference)]. These soluble CD4 proteins advantageously interfere with the $CD4^+$ lymphocyte/HIV interaction by blocking or competitive binding mechanisms which inhibit HIV infection of cells expressing the CD4 protein. The first immunoglobulin-related domain is sufficient to bind gp120/160. By acting as soluble virus receptors, soluble CD4 proteins are useful as antiviral therapeutics to inhibit HIV binding to $CD4^+$ lymphocytes and virally induced syncytia formation.

The CD4 polypeptides useful in this invention include all CD4 polypeptides which bind to or otherwise inhibit gp120/160. These include fragments of CD4 lacking the transmembrane domain, amino acids +371 to +391 of Figures 3A-3B. Such fragments may be truncated forms of CD4 or may be fusion proteins in which the fourth immunoglobulin-related domain is joined directly to the hydrophilic cytoplasmic domain. Because the secondary structure of a polypeptide is important to its function, soluble CD4 proteins preferably will contain enough of a domain to allow an intra-domain disulfide bond but not enough to include the first cysteine of the next immunoglobulin domain. Within this range, certain amino acid sequences bind gp160/120 with greater affinity than others. We shall refer herein to a CD4 polypeptide which includes amino acids

+1 to +X of Figures 3A-3B, and optionally including an N-terminal methionine, as "CD4(X)".

For example, referring now to Figures 3A-3B, a soluble CD4 protein containing the first immunoglobulin-like domain preferably will contain at least amino acids +1 to +84 and at most amino acids +1 to +129. Most preferably, a soluble CD4 protein comprises amino acids +1 to +111 [CD4(111)]. A soluble CD4 protein containing the first two immunoglobulin-like domains preferably will include at least amino acids +1 to +159 and at most amino acids +1 to +302. More preferably, a soluble CD4 protein will include at least amino acids +1 to +175 and at most amino acids +1 to +190. Most preferably, a soluble CD4 protein will include amino acids +1 to +181 [CD4(181)], amino acids +1 to +183 [CD4(183)], or amino acids +1 to +187 [CD4(187)]. A soluble CD4 protein which includes the first four immunoglobulin-like domains preferably will include at least amino acids +1 to +345 and at most amino acids +1 to +375 [CD4(375)]. Any of these molecules may optionally include the CD4 signal sequence, amino acids -25 to -1 of Figures 3A-3B. Also, these molecules may have a methionine residue optionally preceding amino acid +1 of Figures 3A-3B.

Soluble CD4 proteins useful in the fusion polypeptides and methods of this invention may be produced in a variety of ways. We have depicted in Figures 3A-3B the nucleotide sequence of full-length CD4 cDNA obtained from pJOD.sCD4.Y187 and p170.2 and the amino acid sequence deduced therefrom. According to the coordinate system in Figures 3A-3B, the amino terminal amino acid of mature CD4 protein isolated from T cells is lysine, located at nucleotide 136 of Figure 3 [D.R. Littman et al., supra]. Soluble CD4 proteins also include those in which amino acid +62 is

arginine, encoded by CGG, and those in which amino acid +229 is phenylalanine, encoded by TTT. Therefore, when we refer to CD4, we intend to include amino acid sequences including one or both of these substitutions. Soluble CD4 polypeptides may be produced by conventional techniques of oligonucleotide directed mutagenesis and restriction digestion, followed by insertion of linkers, or by digesting full-length CD4 protein with enzymes.

Soluble CD4 proteins include those produced by recombinant techniques according to the processes set forth in copending, commonly assigned United States patent applications Serial No. 094,322, filed September 4, 1987, Serial No. 141,649, filed January 7, 1988 and Serial No. 351,945, filed May 24, 1989 and PCT patent application Serial No. PCT/US88/02940, filed September 1, 1988, and published as PCT patent application WO 89/01940, the disclosures of which are hereby incorporated by reference.

Microorganisms and recombinant DNA molecules characterized by DNA sequences coding for soluble CD4 proteins are exemplified by cultures described in PCT patent application WO 89/01940. They were deposited in the In Vitro International, Inc. culture collection, in Linthicum, Maryland, USA on September 2, 1987 and identified as:

EC100: E.coli JM83/pEC100 - IVI 10146
BG377: E.coli MC1061/pBG377 - IVI 10147
BG380: E.coli MC1061/pBG380 - IVI 10148
BG381: E.coli MC1061/pBG381 - IVI 10149.

Such microorganisms and recombinant DNA molecules are also exemplified by cultures deposited in the In Vitro International, Inc. culture collection on January 6, 1988 and identified as:

BG-391: E.coli MC1061/pBG391 - IVI 10151

BG-392: E.coli MC1061/pBG392 - IVI 10152
BG-393: E.coli MC1061/pBG393 - IVI 10153
BG-394: E.coli MC1061/pBG394 - IVI 10154
BG-396: E.coli MC1061/pBG396 - IVI 10155
203-5 : E.coli SG936/p203-5 - IVI 10156.

Additionally, such microorganisms and recombinant DNA molecules are exemplified by cultures deposited in the In Vitro International, Inc. culture collection on August 24, 1988 and identified as:

211-11: E.coli A89/pBG211-11 - IVI 10183
214-10: E.coli A89/pBG214-10 - IVI 10184
215-7 : E.coli A89/pBG215-7 - IVI 10185.

Multimeric C4bp fusion proteins comprising CD4-C4bp fusion polypeptides are useful in a variety of pharmaceutical compositions and methods. CD4-C4bp fusion proteins advantageously inhibit HIV binding to $T4^+$ lymphocytes by virtue of their competitive binding characteristics. And they actively destroy HIV infected cells expressing the gp120/160 protein and producing HIV. Accordingly, the CD4-C4bp fusion proteins may be used in pharmaceutical compositions and methods to treat humans having AIDS, ARC, HIV infection, or antibodies to HIV. They are also useful to lessen the immuno-compromising effects of HIV infection or to prevent incidence and spread of HIV infection. In addition, these CD4-C4bp fusion proteins and methods may be used for treating AIDS-like diseases caused by retroviruses, such as simian immunodeficiency viruses, in mammals, including humans.

DNA sequences encoding C4bp fusion polypeptides are useful for producing multimeric C4bp fusion proteins. The preferred process involves expressing such DNA sequences in a host that will properly assemble the expressed polypeptides into a multimer.

As is well known in the art, for expression of the DNA sequences of this invention, the DNA sequence should be operatively linked to an expression control sequence in an appropriate expression vector and employed in that expression vector to transform an appropriate unicellular host. Such operative linking of a DNA sequence of this invention to an expression control sequence, of course, includes the provision of a translation start signal in the correct reading frame upstream of the DNA sequence. If a particular DNA sequence being expressed does not begin with a methionine, the start signal will result in an additional amino acid -- methionine -- being located at the N-terminus of the product. While such a methionyl-containing product may be employed directly in the compositions and methods of this invention, it is usually more desirable to remove the methionine before use. Methods are known to those of skill in the art to remove such N-terminal methionines from polypeptides expressed with them. For example, certain hosts and fermentation conditions permit removal of substantially all of the N-terminal methionine in vivo. Other hosts require in vitro removal of the N-terminal methionine. However, such in vivo and in vitro methods are well known in the art.

A wide variety of host/expression vector combinations may be employed in expressing the DNA sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic DNA sequences, such as various known derivatives of SV40 and known bacterial plasmids, e.g., plasmids from E.coli including colE1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, e.g., RP4; phage DNAs, e.g., the numerous derivatives of phage λ,

e.g., NM989, and other DNA phages, e.g., M13 and filamentous single stranded DNA phages; yeast plasmids, such as the 2μ plasmid or derivatives thereof; and vectors derived from combinations of plasmids and phage DNAs, such as plasmids which have been modified to employ phage DNA or other expression control sequences.

In addition, any of a wide variety of expression control sequences -- sequences that control the expression of a DNA sequence when operatively linked to it -- may be used in these vectors to express the DNA sequences of this invention. Such useful expression control sequences, include, for example, the early and late promoters of SV40 or adenovirus, the lac system, the trp system, the TAC or TRC system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α-mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof.

A wide variety of unicellular host cells are also useful in expressing the DNA sequences of this invention. These hosts include well known eukaryotic and prokaryotic hosts, such as strains of E.coli, Pseudomonas, Bacillus, Streptomyces, fungi, such as yeasts, and animal cells, such as CHO and mouse cells, African green monkey cells, such as COS-1, COS-7, BSC 1, BSC 40, and BMT 10, insect cells, and human cells and plant cells in tissue culture. For animal cell expression, we prefer CHO cells and COS-7 cells.

It should of course be understood that not all vectors and expression control sequences will

function equally well to express the DNA sequences of this invention. Neither will all hosts function equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences, and hosts without undue experimentation and without departing from the scope of this invention. For example, in selecting a vector, the host must be considered because the vector must replicate in it. The vector's copy number, the ability to control that copy number, and the expression of any other proteins encoded by the vector, such as antibiotic markers, should also be considered.

In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the system, its controllability and its compatibility with the particular DNA sequence of this invention, particularly as regards potential secondary structures. Unicellular hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of the product coded for on expression by the DNA sequences of this invention to them, their secretion characteristics, their ability to fold proteins correctly, their fermentation requirements and the ease of purification of the products coded on expression by the DNA sequences of this invention.

Within these parameters, one of skill in the art may select various vector/expression control system/host combinations that will express the DNA sequences of this invention on fermentation or in large scale animal culture, e.g., CHO cells or COS-7 cells.

According to one embodiment of this invention, a DNA sequence encoding a CD4-C4bp fusion polypeptide inserted into plasmid pJOD-S (described

herein) and expressed in COS-7 or CHO cells produces fusion polypeptides which naturally assemble into heptameric CD4-C4bp fusion proteins.

The polypeptides and proteins produced on expression of the DNA sequences of this invention may be isolated from fermentation or animal cell cultures and purified using any of a variety of conventional methods. One of skill in the art may select the most appropriate isolation and purification techniques without departing from the scope of this invention.

One can also produce C4bp fusion polypeptides by chemical synthesis using conventional peptide synthesis techniques, such as solid phase synthesis [R.B. Merrifield, "Solid Phase Peptide Synthesis. I. The Synthesis Of A Tetrapeptide", J. Am. Chem. Soc., 83, pp. 2149-54 (1963)]. Multimeric C4bp fusion proteins may then be produced in vitro by forming intra- and inter-C4bp fusion polypeptide disulfide bonds.

Another method useful for producing multimeric C4bp fusion proteins, in addition to genetic fusion and chemical synthesis, is by chemically coupling a functional moiety to the C4bp monomer. This method is useful for both chemical moieties or polypeptide moieties. One may couple the functional moiety to individual C4bp monomers or to C4bp monomers already assembled into a multimer, for example, hC4bp itself or multimeric recombinant hC4bp.

Several methods may be used for chemical coupling. These include, for example, methods using glutaraldehyde [M. Reichlin, "Use of Glutaraldehyde as a Coupling Agent for Proteins and Peptides", Methods In Enzymology, 70, pp. 159-65 (1980)], N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide [T.L. Goodfriend et al., "Antibodies to Bradykinin and Angiotensin: A

Use of Carbodiimides in Immunology", Science, 144, pp. 1344-46 (1964)] or a mixture of N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide and a succinylated carrier [M.H. Klapper and I.M. Klotz, "Acylation with Dicarboxylic Acid Anhydrides", Methods In Enzymology, 25, pp. 531-36 (1972)] or those heterobifunctional or homobifunctional cross-linking agents described in the Pierce Chemical Company Catalog. Since chemical coupling is not limited to one site on the C4bp monomer, it is possible to couple more than one functional moiety to each C4bp monomer. One can also couple the functional moiety to a glycan on the protein using the sodium periodate procedure [P.K. Nakane and A. Kawaoi, "Peroxidase-labeled Antibody: A New Method of Conjugation", J. Histochem. Cytochem., 22, pp. 1084-91 (1984)].

Hetero-multimeric C4bp fusion proteins comprise combinations of different C4bp monomers, different functional moieties, or combinations of both. For example, hetero-multimeric C4bp fusion proteins may comprise combinations of more than one C4bp monomer (i.e., C4bp.SCR4 and C4bp.SCR8) with one type of functional moiety, one type of C4bp monomer with combinations of more than one type of functional moiety (i.e., a recognition molecule and a reporter group), combinations of more than one type of C4bp monomer with combinations of more than one type of functional moiety and combinations in which not all of the C4bp monomers are fused or chemically coupled to functional moieties.

A hetero-multimeric C4bp fusion protein comprising two different polypeptide moieties may advantageously be produced by expressing DNA sequences encoding the two different polypeptides in a single host. Upon expression in an appropriate system, the polypeptides will assemble into multimeric fusion

proteins containing more than one type of functional moiety.

According to an alternate embodiment of this invention, hetero-multimers characterized by polypeptide and chemical moieties, or two different chemical moieties, may also be produced. As described above, a moiety may be chemically coupled to polypeptides before or after assembly.

Hetero-multimeric C4bp fusion proteins are especially useful when the properties of the different moieties complement one another. For example, it is possible to combine receptors that bind to a particular target particle or cell and toxins or anti-retroviral agents in fusion proteins according to this invention to produce targeted toxic or anti-retroviral agents. Polypeptides useful as toxins include, but are not limited to, ricin, abrin, angiogenin, Pseudomonas Exotoxin A, pokeweed antiviral protein, saporin, gelonin and diptheria toxin, or toxic portions thereof. Useful anti-retroviral agents include suramin, azidothymidine (AZT), dideoxycytidine and glucosidase inhibitors such as castanospermine, deoxynojirimycin and derivatives thereof.

Hetero-multimeric C4bp fusion proteins according to this invention are also useful as diagnostic agents to identify the presence of a target molecule in a sample or in vivo. Such proteins comprise one functional moiety which is a recognition molecule, such as an immunoglobulin or a fragment thereof (Fab, dAb) that binds to the target molecule [See Ward et al., supra] and a second functional moiety which is a reporter group, such as a radionuclide, an enzyme (such as horseradish peroxidase) or a fluorescent or chemiluminescent marker. Because multimeric C4bp is large, many reporter groups may be

coupled to it, thereby enhancing the signal. These hetero-multimers may be used, for example, to replace antibodies as recognition molecules that contact the sample in ELISA-type assays, or as in vivo imaging agents.

Hetero-multimeric C4bp fusion proteins according to this invention may also be used as multi-vaccines. For example, such fusion proteins may be constructed using several different antigenic determinants from the same infective agent. Also, one can produce fusion proteins comprising antigenic determinants from several infective agents, such as polio, measles, mumps and others used for childhood vaccination, thus creating a multi-vaccine.

Multimeric C4bp fusion proteins according to this invention also include the normally associated protein S-binding subunit of human C4bp. Such proteins are produced upon transformation of a host with a first DNA sequence encoding a C4bp fusion polypeptide and a second DNA sequence encoding the protein S-binding subunit. Upon expression of these DNA sequences, the C4bp polypeptides will assemble into a multimer associated with the protein S-binding subunit.

It should be understood that while C4bp polypeptides normally assemble into a heptamer (not including the protein S-binding subunit) it is possible that if the monomer polypeptides are either smaller or larger than normal, they may assemble into octamers or hexamers, for example. Therefore, the multimeric C4bp fusion proteins referred to in this application include those other than heptameric C4bp fusion proteins.

The pharmaceutical compositions of this invention typically comprise a pharmaceutically effective amount of a C4bp fusion protein of this invention and a pharmaceutically acceptable carrier.

Therapeutic methods of this invention comprise the step of treating patients in a pharmaceutically acceptable manner with those compositions. These compositions may be used to treat any mammal, including humans.

The pharmaceutical compositions of this invention may be in a variety of forms. These include, for example, solid, semi-solid and liquid dosage forms, such as tablets, pills, powders, liquid solutions or suspensions, liposomes, suppositories, injectable and infusable solutions and sustained release forms. The preferred form depends on the intended mode of administration and therapeutic application. The compositions also preferably include conventional pharmaceutically acceptable carriers and adjuvants which are known to those of skill in the art.

Generally, the pharmaceutical compositions of the present invention may be formulated and administered using methods and compositions similar to those used for pharmaceutically important polypeptides such as, for example, alpha interferon. Thus, the fusion proteins of this invention may be stored in lyophilized form, reconstituted with sterile water just prior to administration, and administered by conventional routes of administration such as parenteral, subcutaneous, intravenous, intramuscular or intralesional routes. An effective dosage may be in the range of about 10-100 μg/kg body weight/day, it being recognized that lower and higher doses may also be useful. It will be understood that conventional doses will vary depending upon the particular molecular moiety involved.

In addition, one may use DNA sequences encoding C4bp fusion polypeptides in somatic gene therapy. This involves, for example, inserting DNA sequences into retroviral-based vectors suitable for

infection of human somatic cells [A. Kasid et al., "Human Gene Transfer: Characterization of Human Tumor-infiltrating Lymphocytes as Vehicles for Retroviral-mediated Gene Transfer", Proc. Natl. Acad. Sci., USA, 87, pp. 473-77 (1990)]. For example, patients with AIDS or ARC could be treated as follows. First, one would prepare a retrovirus characterized by a DNA sequence encoding a CD4-C4bp fusion polypeptide. Then, one would isolate T cells from the patient and infect them, in vitro, with the retrovirus. One would then reintroduce these cells into the patient, where the vector will express and the cell will secrete CD4-C4bp polypeptide.

In order that this invention may be better understood, the following examples are set forth. These examples are for the purposes of illustration only, and are not to be construed as limiting the scope of the invention in any manner.

In the examples that follow, the molecular biology techniques employed, such as cloning, cutting with restriction enzymes, isolating DNA fragments, filling out with Klenow enzyme and deoxyribonucleotides triphosphate (dXTP), ligating, transforming E.coli and the like are conventional protocols exemplified and further described in J. Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

EXAMPLE I -- CLONING OF C4 BINDING PROTEIN

We isolated a cDNA sequence encoding human C4 binding protein in the following manner. We obtained human hepatocytes (Hep G2) from the American Type Culture Collection, Rockville, Maryland, USA, ATCC HB 8065. We isolated polyadenylated mRNA from these cells using the guanidinium isothiocyanate/oligo dT cellulose

procedure [Sambrook et al., supra, pp. 7.19-7.22]. Using 5 μg polyadenylated mRNA, Mo-MLV reverse transcriptase and a primer, C4bp.3, we synthesized antisense C4bp single-stranded cDNA. (The DNA sequences of all oligonucleotide primers and splint probes are given in Figures 5A-5B.) We then synthesized the second strand cDNA using the Gubler-Hoffman technique [U. Gubler and B.J. Hoffman, "A Simple and Very Efficient Method for Generating cDNA Libraries", Gene, 25, pp. 263-69 (1983)].

We amplified cDNA sequences for C4bp using PCR [Sambrook et al., supra, Chapter 14]. We carried out all amplifications using Taq DNA polymerase and primers pre-phosphorylated with T4 polynucleotide kinase and ATP. We used the oligonucleotide C4bp.1 as the sense primer (which hybridized to the antisense strand) and C4bp.2 as the antisense primer. We filled out the amplified fragments with Klenow enzyme and dXTP. This produced a 1746 bp fragment encoding C4bp and bordered by transcriptional start and stop signals. We verified the identity of this fragment by digestion with SnaBI and with PstI. As predicted by the DNA sequence of C4bp, SnaBI digestion produced a 1436 bp fragment and SnaBI/PstI digestion produced a 1047 bp fragment.

Then we inserted the C4bp-encoding fragment into the animal expression vector, pJOD-S, which was created as follows. (See Figures 6B-6C).

First we obtained pJOD-10. As described in European patent application 343,783, pJOD-10 may be prepared as follows. Plasmid pSV2-DHFR, (ATCC 37146, from the American Type Culture Collection) [S. Subramani et al., "Expression of the Mouse Dihydrofolate Reductase Complementary Deoxyribonucleic Acid in Simian Virus 40 Vectors", Molec. Cell. Biol.,

1, pp. 854-64 (1981)] was digested with ApaI and EcoRI and the 4420 bp fragment was isolated. Then, a synthetic double stranded DNA sequence having an ApaI overhang, a DNA sequence encoding nucleotides +190 to +233 of the human gastrin gene [K. Sato et al., "A Specific DNA Sequence Controls Termination of Transcription in the Gastrin Gene", Molec. Cell. Biol., 6, pp. 1032-43 (1986) Figure 4], an XhoI site, and an EcoRI overhang was produced. This oligonucleotide was ligated with the 4420 bp fragment obtained from pSV2-DHFR and the resulting plasmid was called pDT4. Plasmid pDT4 was then digested with AatII and XhoI and the 4391 bp fragment was isolated. The Mullerian inhibiting substance expression vector pD1 [R.L. Cate et al., "Isolation of the Bovine and Human Genes for Mullerian Inhibiting Substance and Expression of the Human Gene in Animal Cells", Cell, 45, pp. 685-96 (1986)] was then digested with AatII and SalI and the resulting 5462 bp fragment was isolated. This fragment was ligated with the 4391 bp fragment pDT4 to produce pJOD-10.

We digested pJOD-10 with HindIII and BstEII and isolated the large fragment which did not encode Mullerian inhibiting substance. We blunt-ended the fragment, ligated SalI linkers to the ends and self-ligated the vector. This produced pJOD-S.

We then prepared pJOD-S for insertion of the C4bp-encoding fragment. We linearized the plasmid at the unique SalI site by digestion with SalI and filled it out with Klenow enzyme and dXTP. We then ligated the C4bp-encoding fragment to it using T4 DNA ligase and ATP. We introduced the ligation mixture into E.coli HB101 by electroporation. We performed electroporation at 25 μFD, 2.5 kV, 200 ohms using a BioRad GENE PULSER® according to the protocol supplied

with the instrument [Biorad Catalog, "Bacterial Electro-transformation and Pulse Controller Instrument Manual", #165-2098, version 2-89]. Then we identified plasmids containing the insert in the proper orientation by hybridizing with ^{32}P-labelled synthetic oligonucleotide splint probes C4bp.9 and C4bp.10. Splint probes are 30 base long synthetic oligonucleotides that hybridize across the point of fusion between an insert and a vector. We called the resulting plasmid pJOD.C4bp. We have deposited one isolate of this plasmid, pJOD.C4bp.3.

EXAMPLE II -- EXPRESSION OF A C4bp HEPTAMER

We introduced supercoiled plasmid DNA from five isolates of pJOD.C4bp into COS-7 cells by electroporation to test them for expression of recombinant human C4bp (rhC4bp). We performed electroporation at 280 V and 960 μFD using 1 x 10^7 cells in 800 μl of 20 mM HEPES, pH 7.05, 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4 and 6 mM dextrose with 20 μg supercoiled plasmid and 380 μg sonicated salmon sperm DNA. As a control, we used plasmids containing the C4bp fragment inserted backwards into the vector, so that they would not produce C4bp at all.

We plated the transformed cells in 100 mm dishes or T75 tissue culture flasks in DMEM medium containing 10% FBS, 4 mM glutamine, 20 mM HEPES (pH 6.8) at 37°C and 5% CO_2. We then assayed the culture fluid after seventy-two hours for rhC4bp and characterized the product using three methods: immunoprecipitation, gel filtration and immunodetection on Western blots. In these assays, we compared the rhC4bp produced to the naturally occurring form in human serum, hC4bp. Our results indicated that the transfected COS-7 cells produced a heptameric C4bp

protein with properly folded SCRs resembling naturally occurring hC4bp in molecular weight and possessing epitopes recognized by anti-hC4bp antisera. Recombinant human C4bp differs from the serum form in that it lacks a protein S-binding subunit.

A. Immunoprecipitation of rhC4bp

In the first assay, we immunoprecipitated both rhC4bp and hC4bp using two different hC4bp-specific antisera and compared the size of the precipitated proteins before and after reduction of disulfide bonds by means of standard sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

A comparison of rhC4bp and hC4bp isolated by immunoprecipitation showed that rhC4bp produced in COS-7 cells assembled properly into a heptameric C4bp protein via disulfide bonds.

We performed immunoprecipitation on human serum (Gibco, Grand Island, New York) and COS-7 culture fluid as follows. The sample had been pretreated by adding to 1 ml of serum a 0.05 ml suspension of immobilized Protein G (Gibco, Rockford, Illinois) and agitating the mixture for 30 minutes at room temperature. Then we pelleted the Protein G particles in a centrifuge and used the supernatant for the immunoprecipitation. We incubated the supernatant with 0.05 ml of either polyclonal rabbit anti-hC4bp antiserum (Calbiochem Corp., San Diego/La Jolla, California) or polyclonal sheep anti-hC4bp antiserum (Biodesign International, Kennebunkport, Maine) and incubated for 1 hour at room temperature. We then added 0.05 ml of Protein G suspension and incubated for 1 hour at room temperature. The pellet was centrifuged and resuspended in 50 mM tris hydroxy amino-methane (pH 8.0) (Tris, Sigma Chemical Corp., St. Louis, Missouri)

containing 100 mM NaCl and 1% Tween 20 (Pierce). We centrifuged the solution again and removed the supernatant. We repeated this procedure twice with the same buffer and with 10 mM Tris, pH 7.4. Finally, we resuspended the pellet in 0.15 ml standard Laemmli sample buffer and heated the solution in a boiling water bath for 5 minutes. Then we determined the molecular weights of the precipitated proteins by 5% or 12% SDS-PAGE.

Human serum C4bp produced a band of about 530 kD, representing a C4bp heptamer bound to the 45 kD protein S-binding subunit. The rhC4bp produced a band of about 490 kD, the predicted molecular weight of a heptameric C4bp protein not including the protein S-binding subunit. The control sample, with the DNA insert in the non-expressing orientation, did not produce a band.

To verify that the precipitated rhC4bp was a heptamer, we performed SDS-PAGE on the immunoprecipitated proteins using the reducing agent, 2-mercaptoethanol, which breaks disulfide-bonded proteins into their polypeptide subunits. We found that both hC4bp and rhC4bp produced bands of 70 kD, the size of C4bp polypeptide.

We also carried out immunoprecipitation on rhC4bp which had been expressed in COS-7 cells in the presence of ^{35}S-labelled cysteine (New England Nuclear, Boston, Massachusetts). We precipitated the resulting ^{35}S-labelled protein using the above mentioned rabbit anti-hC4bp serum and analyzed it on 4%-20% gradient SDS-PAGE. Under non-reducing conditions we detected on an autoradiograph a high molecular weight band equivalent to the above-described 490 kD protein. After reduction, this band disappeared and gave rise to a band of about 70 kd. Both bands were absent in the

negative control sample. This confirmed that the C4bp polypeptides were assembling into a heptamer.

B. Immunodetection of rC4bp on Western blot

We next precipitated rhC4bp and hC4bp unselectively with trichloroacetic acid (TCA) and identified the proteins on Western blot using three different hC4bp-specific antisera. Immunodetection of rhC4bp and hC4bp on Western blot under non-reducing and reducing conditions suggested that COS-7 cells produced a rhC4bp with properly folded disulfide bonds.

We precipitated rhC4bp from 10-fold concentrated cell culture fluid (concentrated via ultrafiltration, CENTRIPREP 30®, Amicon) and hC4bp from serum by addition of 12% w/v TCA. After one hour in the ice bath, the proteins were pelleted by centrifugation in an EPPENDORF® centrifuge (Eppendorf) at 10,000 g for 10 min at 4°C. The pellet was resuspended in 1 ml of cold acetone (-20°C) and was immediately repelleted under the above conditions. We repeated this wash step once. Finally, we dissolved the protein pellet in standard Laemmli sample buffer and heated the solution in a boiling water bath for 5 minutes. We then separated the proteins on a 5% acrylamide gel (SDS-PAGE) under non-reducing and reducing conditions (using 2-mercaptoethanol). We transferred these proteins on nitrocellulose using standard immunoblotting techniques. Then we examined the blots by immunodetection.

We blocked non-specific binding on these blots by incubating them in Dulbecco's PBS with 5% non-fat dry milk (Carnation, Los Angeles, California). Next, we incubated the blot in a 5% milk solution containing a primary antibody at a dilution of 1:500. As primary antibody, we used polyclonal rabbit anti-

hC4bp antiserum (Calbiochem), polyclonal sheep anti-hC4bp antiserum (Biodesign) or a monoclonal murine anti-hC4bp antibody (Quidel, San Diego, California). After one hour at room temperature, we washed the blot three times with 0.05% Tween-20 in PBS. Then we incubated the blot with the secondary antibody at a dilution of 1:1000 in the 5% milk solution. As secondary antibody, we used commercially available horseradish peroxidase conjugated antibodies directed to rabbit IgG, sheep IgG or mouse IgG, respectively (Amersham Corp., Arlington Heights, Illinois or Calbiochem). After an incubation period of one hour at room temperature, we repeated the above described wash procedure. We visualized antibody-antigen complexes by incubation with the horseradish peroxidase substrates 4-chloro-1-naphthol (0.02%, w/v, Sigma) and hydrogen peroxide (0.03%, Sigma) in PBS.

In the case of rhC4bp, we detected a single band at a molecular weight of ca. 490 kD. In the case of the human serum, the band had been shifted to slightly higher molecular weight (ca. 530 kD) which may reflect the presence of the additional 45 kD protein S binding subunit. We detected no proteins in a negative control. These results again demonstrated that the recombinant form of rhC4bp had assembled into a heptamer which was recognized by anti-hC4bp antisera and which resembled in molecular weight the naturally occurring form minus the protein S binding subunit. After reduction and separation of the proteins on a 12% acrylamide gel (SDS-PAGE), we did not detect any protein in any case using the above antisera. Thus, the antisera used had only recognized hC4bp having correctly folded disulfide bridges. This provided additional evidence that the recombinant form of hC4bp which had been well detected with the antisera under

nonreducing conditions, was similar or identical to the naturally occurring form of that protein without the protein S binding subunit.

C. Gel Permeation Chromatography and Identification of hC4bp- and rhC4bp-containing Fractions via a hC4bp-specific ELISA

We also separated rhC4bp and hC4bp according to size, under conditions that preserved the native structure of the protein, using a high performance liquid chromatography (HPLC) gel permeation technique. Subsequently, we identified the peak positions and thus the approximate molecular weight of the proteins by subjecting fractions to a hC4bp-specific enzyme linked immunosorbent assay (ELISA).

HPLC of rC4bp and hC4bp supported our previous results, indicating that COS-7 cells produced rC4bp in heptameric form. It also suggested that hC4bp exists bound to a natural ligand, such as protein S or C4b (a fragment of C4).

We equilibrated a TSK-4000SWXL® HPLC-gel permeation column (300 x 7.8 mm, Toyo Soda, TOSOM Corp., Japan) in 50 mM phosphate buffer, pH 7.0, 100 mM sodium chloride. We calibrated the column using the molecular weight markers thyroglobulin (ca. 670 kD), ferritin (ca. 440 kD) and catalase (ca. 230 kD) (all from Pharmacia-LKB) and detected the eluted peaks using a UV detector at 280 nm. We loaded human serum (0.05 ml diluted to 0.5 ml with PBS) or a 10-fold concentrated cell culture supernatant (0.5 ml) containing the expressed rhC4bp onto this column. We eluted the proteins from the column with equilibration buffer at a linear flow rate of 1 ml/min, collecting 0.5 ml fractions.

We assayed the collected fractions using a hC4bp-specific sandwich ELISA (ELISA 9). (See Figure 10 for a summary of the ELISA assays described in this specification.) More specifically, we coated a 96 well standard ELISA plate (Immulon II, Dynatech Laboratories) with a polyclonal sheep anti-hC4bp antibody (Biodesign) at a protein concentration of 0.005 μg/ml in 0.05 M bicarbonate buffer, pH 9.0. We incubated the plate at 4°C overnight. We blocked the non-specific sites in the wells by the addition of 2% milk in PBS and incubated it for at least 30 minutes at room temperature. We washed the ELISA plate three times with 0.05% Tween-20 in PBS and added 0.05 ml aliquots of each fraction from the HPLC-column diluted 1:2 or 1:10 in PBS containing 2% nonfat dry milk. Then we incubated the plates for 3 hours at room temperature and washed them as described above. We added 0.05 ml of a polyclonal rabbit anti-hC4bp serum (Calbiochem), optimally diluted (1:3000) in PBS containing 2% milk, to each well. We incubated the plate for 1 hour at room temperature and then washed it as above. Finally, we added 0.05 ml horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (Organon Teknika Corp., West Chester, Pennsylvania), appropriately diluted in PBS containing 2% milk and incubated the plate for 1 hour at room temperature. After washing the plate we added OPD* (o-phenylenediamine) (Calbiochem) as the substrate for HRP in order to detect colorimetrically the amount of HRP-antibody bound. We quantitated the results by reading the absorbance of each well at 490 nm.

---

* This is a potential carcinogen which should be detoxified before disposal using a solution of: 50g K_2CrO_7, 25 ml ION H_2SO_4, 145 ml H_2O.

We then plotted the results against the fraction numbers as shown in Figure 8. As demonstrated in that figure, the elution profile of rhC4bp peaked at about 490 kD, consistent with other results that indicate that the rhC4bp which had been expressed in COS-7 cells was present as a heptamer exposing correctly folded immune epitopes. However, the elution profile of hC4bp peaked somewhat higher than 700 kD, in contrast to the expected 530 kD. This may have resulted from non-covalent binding of the molecule to its natural ligands, C4 and protein S. We have also observed the presence of protein S in these fractions (data not shown).

EXAMPLE III -- A VECTOR EXPRESSING rsCD4(187)

We next produced pJOD.sCD4.Y187.SnaB1, a plasmid characterized by a DNA sequence which encodes rsCD4(187). We began with pJODsCD4, described in PCT application WO 89/01949, which may be prepared as follows (see Figure 6G). Plasmid pBG391 (IVI 10149) was digested with BamHI and BglII and filled out with Klenow enzyme and dXTP. Double stranded XhoI linkers, having the sequence 5' CCTCGAGG were ligated to the ends with T4 DNA ligase. The mixture was digested with XhoI and the 1407 bp fragment encoding rsCD4 was isolated. Then pJOD-S, described in Example I, was digested with SalI and phosphorylated with alkaline phosphatase. The resulting 1407 bp fragment was ligated into the SalI-digested pJOD-S with T4 DNA ligase. This produced pJODsCD4.

We then modified pJODsCD4 to create a SnaBI cloning site (see Figure 6H). We introduced the SnaBI cleavage site directly after the sequence encoding the tyrosine residue at position 187 of CD4 as follows. We created a NheI/BglII linker containing the unique SnaBI

site by hybridizing prephosphorylated C4bp.7 with prephosphorylated C4bp.8. We then digested pJODsCD4 completely with NheI and partially with BglII. We added the linker to this digestion mixture and ligated the fragments with T4 DNA ligase. Then we electroporated the ligation mixture into E.coli HB101 using the method of Example I. We identified plasmids containing the synthetic fragment bordering the correct BglII site of the large digestion fragment by hybridization with the ^{32}P-labelled splint probe C4bp.11. We called such plasmids pJOD.sCD4.Y187.SnaB1.

EXAMPLE IV -- CLONING CD4-C4bp FUSION POLYPEPTIDES

We next constructed three clones that expressed CD4-C4bp fusion proteins (see Figure 7). This involved inserting a DNA sequence encoding C4bp, or a fragment thereof, into the SnaBI site of pJOD.sCD4.Y187.SnaB1. We used three DNA sequences encoding, respectively, SCR8-SCR1 and the C4bp core, SCR4-SCR1 and the C4bp core, or SCR1 and the C4bp core. We generated the DNA sequences by PCR, as follows.

We produced a 1648 bp fragment of C4bp encoding SCR8-SCR1 and the C4bp core by performing PCR with pJOD.C4bp.3 linearized with NotI. We used the prephosphorylated sense primer SCR.8 and the prephosphorylated antisense primer C4bp.2 (Figures 5A-5B). We then repaired the ends of the PCR products with Klenow enzyme and dXTP. We ligated the resulting fragment with pJOD.sCD4.Y187.SnaB1 which had been previously linearized with SnaBI. We electroporated the ligation mixture into E.coli HB101. Then we identified plasmids in which the tyrosine-encoding sequence bordered SCR8 by hybridization to the ^{32}P-labelled oligonucleotide splint probe C4bp.12. We named such plasmids pJOD.sCD4.Y187.SCR8. We called the fusion polypeptide

expressed by such plasmids, CD4(187)-C4bp(SCR8). We obtained isolates of those plasmids, pJOD.sCD4.Y187.SCR8.2 and pJOD.sCD4.Y187.SCR8.3.

We produced a 1089 bp fragment of C4bp encoding SCR5-SCR1 and the C4bp core as above, except that we used the sense primer 312.21 instead of SCR.8. Then we ligated the fragment with pJOD.sCD4.Y187.SnaB1, which had been previously linearized with SnaBI. We electroporated the ligation mixture into E.coli HB101. Then we identified plasmids in which the tyrosine-encoding sequence bordered SCR5 by hybridization to the ^{32}P-labelled oligonucleotide splint probe 312.36. We named such plasmids pJOD.sCD4.Y187.SCR5. We called the fusion polypeptide expressed by such plasmids, CD4(187)-C4bp(SCR5). We obtained two isolates of those plasmids, pJOD.sCD4.Y187.SCR5.1 and pJOD.sCD4.Y187.SCR5.2.

We produced a 908 bp fragment of C4bp encoding SCR4-SCR1 and the C4bp core as above, except that we used the sense primer SCR.4 instead of SCR.8. Then we ligated the fragment with pJOD.sCD4.Y187.SnaB1, which had been previously linearized with SnaBI. We electroporated the ligation mixture into E.coli HB101. Then we identified plasmids in which the tyrosine-encoding sequence bordered SCR4 by hybridization to the ^{32}P-labelled oligonucleotide splint probe C4bp.13. We named such plasmids pJOD.sCD4.Y187.SCR4. We called the fusion polypeptide expressed by such plasmids, CD4(187)-C4bp(SCR4). We obtained two isolates of those plasmids, pJOD.sCD4.Y187.SCR4.2 and pJOD.sCD4.Y187.SCR4.3.

We produced a 711 bp fragment of C4bp encoding SCR3-SCR1 and the C4bp core as above, except that we used the sense primer 312.20 instead of SCR.8. Then we ligated the fragment with pJOD.sCD4.Y187.SnaB1,

which had been previously linearized with SnaBI. We electroporated the ligation mixture into E.coli HB101. Then we identified plasmids in which the tyrosine-encoding sequence bordered SCR3 by hybridization to the ^{32}P-labelled oligonucleotide splint probe 312.35. We named such plasmids pJOD.sCD4.Y187.SCR3. We called the fusion polypeptide expressed by such plasmids, CD4(187)-C4bp(SCR3). We obtained two isolates of those plasmids, pJOD.sCD4.Y187.SCR3.2 and pJOD.sCD4.Y187.SCR3.3.

We produced a 353 bp fragment of C4bp encoding SCR1 and the C4bp core as above, except that we used the sense primer SCR.1, instead of SCR.8. Then we ligated the fragment with pJOD.sCD4.Y187.SnaB1, which had been previously linearized with SnaBI. We electroporated the ligation mixture into E.coli HB101. Then we identified plasmids in which the tyrosine-encoding sequence bordered SCR1 by hybridization to the ^{32}P-labelled oligonucleotide splint probe C4bp.17. We named such plasmids pJOD.sCD4.Y187.SCR1. We called the fusion polypeptide expressed by such plasmids, CD4(187)-C4bp(SCR1). We obtained three isolates of those plasmids, pJOD.sCD4.Y187.SCR1.1, pJOD.sCD4.Y187.SCR1.2 and pJOD.sCD4.Y187.SCR1.3.

EXAMPLE V -- EXPRESSION AND PURIFICATION OF CD4-C4bp MULTIMERIC FUSION PROTEINS

We transformed COS-7 cells, as previously described, with all the isolates described in Example IV and tested the cultures for the expression of CD4-C4bp multimeric fusion proteins. Specifically, we electroporated supercoiled plasmid DNA from each of the isolates into COS-7 cells. After 72 hours of expression, culture fluids were assayed by several

different ELISAs. Our results showed that the cells expressed CD4-C4bp multimeric fusion proteins.

A. Expression of CD4(187)-C4bp(SCR4)

We tested the conditioned medium of transformed cells for the expression of CD4(187)-C4bp(SCR4) by both immunodetection on Western blots and by ELISA assay. The results indicated that these cells expressed CD4(187)-C4bp(SCR4) and that the polypeptides expressed had assembled into a heptamer.

1. ELISA Assays

We performed two ELISA assays that together demonstrated the production of a multimeric CD4(187)-C4bp(SCR4) polypeptide. In all ELISA assays described herein, we washed the plates between steps with 0.05% TWEEN 20 in PBS.

We performed the two ELISA assays as follows. We coated Immulon II plates with either rabbit anti-hC4bp (ELISA 1) or the anti-CD4 monoclonal 6C6 (a gift of Biogen, Inc., Cambridge, Massachusetts) (ELISA 2) by adding 50 µl/well of a 5 µg/ml solution of either antibody in sodium bicarbonate buffer pH 9.0 and incubating the plates overnight at 4°C. After removing the coating solution, we blocked non-specific binding by adding 200 µl/well of 2% non-fat dry milk in PBS and incubating for at least 30 minutes at room temperature. Then, we removed the blocking solution and added 50 µl/well of sample (conditioned medium or conditioned medium diluted in 2% non-fat dry milk) in PBS and incubated for three hours at room temperature. We removed this liquid and added 50 µl/well of the detector antibody, optimally diluted (1:1000) HRP-conjugated 6C6. 6C6 is an anti-human CD4 murine monoclonal that blocks CD4-gp120 binding. (Another

antibody that blocks CD4 binding to gp120, which may be used in place of 6C6, is anti-Leu-3a, available from Becton Dickinson, Mountain View, California). We incubated the plates for 1 hour at room temperature. We removed this solution and added OPD. We again incubated the plates for 20 to 30 minutes at room temperature and stopped the color reaction with 1N sulfuric acid. Then we measured the O.D. at 490 nm.

Both assays gave positive results. The assay in which we coated the plate with anti-hC4bp antibody detected the presence of polypeptides which contained both hC4bp and CD4 epitopes. The assay in which we coated the plate with the 6C6 monoclonal confirmed the presence of multimers because only multimers possess multiple binding sites able to bind simultaneously to more than one copy of 6C6.

2. Immunodetection on Western blot

Immunodetection of the conditioned-media on Western blot confirmed that the transformed cells had produced CD4(187)-C4bp(SCR4) and further indicated that the polypeptides had assembled into heptamers. We carried out immunoprecipitation on two samples of conditioned media as described in Example II. On the first sample, we used a polyclonal rabbit anti-hC4bp antiserum (Calbiochem). On the second sample, we used 6C6 antibody. We separated the immunoprecipitated proteins in each sample on 5% SDS-PAGE and then transferred the proteins onto nitrocellulose under standard electroblotting conditions. We probed the resulting Western blots with each of two antisera. The first was polyclonal sheep anti-hC4bp antiserum (Biodesign). The second was polyclonal anti-hCD4 antiserum (a gift of Biogen, Inc.). We carried out immunodetection as described in Example I. Both the

anti-hC4bp antiserum and the anti-hCD4 monoclonal detected a protein on both blots with the expected molecular weight of the heptameric form of CD4(187)-C4bp(SCR4), i.e., 400 kD - 500 kD.

We performed the same immunodetection procedure using controls -- rhC4bp and human serum. In this case, anti-hC4bp also detected a high molecular weight form of protein. However, the anti-hCD4 monoclonal failed to detect any protein in the control samples.

The fact that the same protein was precipitated from conditioned medium by both the anti-hCD4 and the anti-hC4bp antiserum demonstrated that CD4(187)-C4bp(SCR4) actually had been expressed as a fusion polypeptide and assembled to a heptameric form.

We also carried out an immunoprecipitation on CD4(187)-C4bp(SCR4) which had been expressed in COS-7 cells in the presence of ^{35}S-labelled cysteine (New England Nuclear). We precipitated the resulting ^{35}S-labelled protein using the above mentioned rabbit anti-hC4bp serum and analyzed it on 4%-20% gradient SDS-PAGE. Under non-reducing conditions, we detected on an autoradiograph a high molecular weight band at 400 kD - 500 kD. After reduction, this band disappeared and gave rise to a band of 53.5 kD. Both bands were absent in the negative control sample. This confirmed our previous results that CD4(187)-C4bp(SCR4) had been expressed and assembled into a heptamer.

B. Purification of Multimeric CD4(187)-C4bp(SCR4) By Column Chromatography

We purified CD4(187)-C4bp(SCR4) using conventional column chromatography techniques. We collected 20 l of conditioned medium derived from a transformed CHO cell line that stably expressed

CD4(187)-C4bp(SCR4). We prepared the transformed CHO cell line by transforming CHO cells with pJOD.sCD4.Y187.SCR4.2 by means of electroporation, as described in U.S. patent 4,956,288 to Barsoum. Preferably, we purified the CD4-C4bp fusion protein using three columns sequentially: first, a FAST S® (Pharmacia) ion-exchange column; second, a Cu chelate SEPHAROSE® (Pharmacia) column; and third, a Zn chelate SEPHAROSE® (Pharmacia) column. However, the fusion protein may be partially purified on FAST S® alone.

To perform FAST S® chromatography, we adjusted the conditioned medium to pH 8.0 with sodium hydroxide and filtered it through a 5 μm prefilter and a 0.45 μm filter. We loaded the filtered medium on a 300 ml (ID 50 mm) FAST Q® ion-exchange column at 3 ml/cm^2hr. We washed the column with 5 column volumes of 50 mM HEPES buffer, pH 8.0, containing 200 mM NaCl. We then eluted the CD4-C4bp fusion protein with 50 mM HEPES, pH 8.0, containing 250 mM NaCl.

We further purified the eluate using Cu and Zn chelate columns (40 ml, ID 25 mm) sequentially. We carried out all procedures on these columns at 4°C and at a flow rate of 0.4 ml/cm^2hr. Each wash, as we describe, was carried out with 80 ml of buffer.

We prepared the Cu chelate column by loading chelating SEPHAROSE® with Cu ions using an aqueous 50 mM $CuCl_2$ solution. We washed the column twice, first with 500 mM Tris, 500 mM NaCl, pH 8.0, followed by with 10 mM Tris, 500 mM NaCl, pH 8.0.

Then we loaded the eluate from the FAST Q® column on the Cu chelate column. We washed the column three times, first with 500 mM Tris, 500 mM NaCl, pH 8.0; second, with 10 mM Tris, pH 8.0; and third, with 10 mM Tris, 100 mM imidazole, pH 8.0. Then we eluted the CD4-C4bp fusion protein with 10 mM Tris, 50 mM

EDTA, pH 8.0. The eluate was dialyzed into 10 mM Tris, 500 mM NaCl, pH 8.0.

Then we prepared a Zn chelate SEPHAROSE® column similarly to the Cu chelate column but used $ZnSO_4$ instead of $CuCl_2$.

We loaded the dialyzed eluate from the Cu chelate column onto the Zn chelate column. After the sample was loaded, we washed the column three times, as before, first with 500 mM Tris, 500 mM NaCl, pH 8.0; second with 10 mM Tris, pH 8.0, and third with 10 mM Tris, 100 mM imidazole, pH 8.0. Then we eluted the CD4-C4bp fusion protein with 10 mM Tris, 25 mM imidazole, 500 mM NaCl, pH 8.0.

We stored the protein in 20 mM HEPES, pH 8.0, containing 500 mM NaCl. This procedure resulted in a significant concentration and purification of CD4(187)-C4bp(SCR4), as determined by SDS-PAGE.

We then examined the purified CD4(187)-C4bp(SCR4) by electron microscopy. To do this, we mixed the CD4-C4bp fusion protein with glycerol, sprayed it on a carbon-coated grid and rotary shadowed it with platinum using conventional techniques. At high magnification, the molecule appeared to have a "spider-like" shape, with multiple rod-like arms extending from its center.

C. Affinity Purification of CD4(187)-C4bp(4SCR)

As an alternative to the column chromatography purification described above, we also purified CD4(187)-C4bp(SCR4) as follows. We concentrated conditioned media from CHO cells producing the CD4(187)-C4bp(SCR4) multimeric protein 50-60 fold at 4°C using a S10Y30® spiral cartridge (Amicon, Danvers, Massachusetts). We passed the concentrated media through a 1D7-CNBr SEPHAROSE® affinity column

equilibrated in PBS, calcium and magnesium free, at approximately one column volume/hour at 4°C. 1D7 is a monoclonal antibody that binds to the second immunoglobulin-related domain of CD4 (1D7 was a gift from Patricia Chisholm of Biogen, Inc.). It was produced by immunizing a mouse with transfected CHO cells that expressed full-length CD4 protein. We have deposited a hybridoma line that produces 1D7, designated Monoclonal Antibody 1D7.G11, with the In Vitro International, Inc. culture collection. CNBr SEPHAROSE® was purchased from Sigma Chemical Corp., St. Louis, Missouri.) 1D7 was coupled to the resin at a density of 0.5 mg/ml, essentially following the manufacture's instructions. Generally, the antibody concentration on the resin should be kept as close to 0.5 mg/ml as possible to achieve maximum binding and elution of the fusion protein.

We washed the loaded column with 3-5 column volumes of PBS, followed by PBS with 0.5 M NaCl and PBS again. We eluted the bound protein with 20 mM triethylamine pH 11.5 (Pierce, Rockford, Illinois), 0.5 M NaCl. We immediately neutralized the fractions with 1/50 volume of 1M HEPES pH 6.8 and stored them at 4°C. SDS-PAGE analysis revealed substantially pure CD4(187)-C4bp(SCR4) protein. To exchange the buffer to PBS and remove the remaining impurities, we concentrated the preparations to about 2 mg/ml by ultrafiltration on a YM30® membrane (Amicon, Danvers, Massachusetts). We applied the concentrate to a 1.6 x 50 cm SUPEROSE-6® size exclusion column (preparative grade, Pharmacia/LKB, Piscataway, New Jersey) equilibrated in PBS. The fractions containing the fusion protein were identified by SDS-PAGE and Coomassie stain, and pooled and sterile filtered through a 0.22 μ MILLEX-GV® filter (Millipore, Bedford, Massachusetts). By amino acid

analysis, we determined the extinction coefficient at 280 nm of CD4(187)-C4bp(SCR4) to be 1.44 A_{280} units/mg-ml in a 1 cm light path cuvette. We stored the final material at 4°C until use. For maximum activity, the fusion protein should be used within 3-5 days after purification.

EXAMPLE VI -- BIOLOGICAL ACTIVITY OF CD4-C4bp FUSION PROTEINS

A. Binding of CD4(187)-C4bp(SCR4) to gp120

We demonstrated the ability of CD4(187)-C4bp(SCR4) to bind to gp120 of the HIV virus by means of two types of assays: ELISA assays and a syncytia blocking assay.

1. ELISA assay

We performed an ELISA assay (ELISA 3) that demonstrated the ability of CD4(187)-C4bp(SCR4) to bind gp120. We coated Immulon II plates with gp120 by adding 50 µl/well of a 5 µg/ml solution of gp120 (commercially available from American Bio-Technologies, Inc., Cambridge, Massachusetts) in PBS and incubating the plates overnight at 4°C. After removing the coating solution, we blocked non-specific binding by adding 200 µl/well of 2% non-fat dry milk in PBS and incubated the plates for at least 30 minutes at room temperature. Then we removed the blocking solution and added 50 µl/well of sample (conditioned medium or conditioned medium diluted in 2% non-fat dry milk in PBS) and incubated the plates for three hours at room temperature. We then removed the liquid and added the detector antibody, rabbit anti-hC4bp (Calbiochem) optimally diluted (1:3333). We again incubated the plates for 1 hour at room temperature. Then we removed this solution and added 50 µl/well of

optimally diluted (1:1000) HRP-conjugated goat anti-rabbit-IgG (Organon Teknika). We incubated the plates for 1 hour at room temperature, then removed the solution, added OPD and again incubated the plates for 20 to 30 minutes at room temperature. We stopped the color reaction with 1N sulfuric acid. We measured O.D. at 490 nm. This assay gave positive results, indicating that the CD4(187)-C4bp(SCR4) fusion polypeptide bound to gp120.

We performed another assay (ELISA 4) confirming these results and also demonstrating that CD4(187)-C4bp(SCR4) had assembled into a multimer. We coated Immulon II plates with gp120 by adding 50 µl/well of a 5 µg/ml solution of gp120 in PBS and incubating the plates overnight at 4°C. After removing the coating solution, we blocked non-specific binding by adding 200 µl/well of 2% non-fat dry milk in PBS and incubating for at least 30 minutes at room temperature. then we removed the blocking solution and added 50 µl/well of sample (conditioned medium or conditioned medium diluted in 2% non-fat dry milk in PBS) and incubated for three hours at room temperature. We removed this liquid, added 50 µl/well of optimally diluted (1:1000) HRP-conjugated 6C6 and incubated the plates for 1 hour at room temperature. We removed this solution and added OPD. We incubated for 20 to 30 minutes at room temperature and stopped the color reaction with 1N sulfuric acid. We measured the O.D. at 490 nm. This second assay confirmed the presence of multimers because only multimers possess multiple binding sites capable of binding simultaneously to gp120 and the 6C6 monoclonal.

2. Syncytia blocking assay

HIV-infected cells, which express gp120 on their surface, fuse with CD4-expressing cells to form multinucleate cells (syncytia). Molecules that bind to gp120 tend to block the formation of syncytia.

We carried out a C8166 cell fusion assay as described in B.D. Walker et al., "Inhibition of Human Immunodeficiency Virus Syncytium Formation and Virus Replication by Castanospermine", Proc. Natl. Acad. Sci. USA, 84, pp. 8120-24 (1987). We incubated 5 x 10^3 H9 cells chronically infected with HTLV-IIIB in 100 μl RPMI 1640 medium containing 10 mM HEPES, pH 6.8, 2 mM glutamine and supplemented with 20% fetal bovine serum for 30 minutes at 37°C in 5% CO_2 with various concentrations of CD4(187)-C4bp(SCR4). (H9 cells are available from the AIDS Research and Reference Reagent Program, NIH, Bethesda, Maryland.) We then added 15 x 10^3 C8166 cells (a $CD4^+$ transformed human umbilical cord blood lymphocyte line) [J. Sodroski et al., "Role of HTLV-III/LAV Envelope in Syncytium Formulation and Cytopathicity", Nature, 322, pp. 470-74 (1986)], in 100 μl media to a final volume of 200 μl in each well and incubated at 37°C in 5% CO_2. (C8166 cells were the gift of Dr. Robert Schooley, Massachusetts General Hospital, Boston, Massachusetts). We then counted total number of syncytia per well at 2 hours and 4 hours after adding the C8166 cells. Parallel co-cultivations used transient fluid from COS-7 cells transfected with pJOD.C4bp.3 (negative control) or OKT4A at 25 μg/ml (positive control). (OKT4A is available from Ortho Diagnostics Systems, Raritan, New Jersey). We considered a positive result as a 50% reduction in syncytia compared to controls. While fluid from the cells transfected with pJOD.C4bp.3 did not inhibit syncytia formation, fluid from cells

transfected with OKT4A and with CD4(187)-C4bp(SCR4) significantly inhibited syncytia formation. This indicated that the fusion protein bound to gp120/160 on the H9 cell surface.

B. CD4(187)-C4bp(SCR4) Blocks Replication of HIV-1 In Vitro

We tested the ability of CD4(187)-C4bp(SCR4) to block HIV-1 replication in vitro in a microreplication assay, essentially as described in M. Robert-Guroff et al., Nature, 316, pp. 72-74 (1985), however we performed the incubation at 37°C rather than 4°C as described therein.

More specifically, we preincubated HIV-1 (20 µl; 100 $TCID_{50}$) (prepared as described in part C, below) and C8166 cells (10 µl; 40,000 cells) with or without 20 µl aliquots of serial dilutions of CD4(187)-C4bp(SCR4) or recombinant soluble CD4 protein (RECEPTIN® brand rsCD4 was the gift of Biogen, Inc.). The rsCD4 that we used was derived from a Chinese hamster ovary cell line transformed with pBG391 (Example III, supra.) In these assays, we used two different preparations of CD4(187)-C4bp(SCR4). One preparation was conditioned cell culture fluid from a stable CHO cell line (Example V, section B, infra) which synthesizes and secretes CD4(187)-C4bp(SCR4). The other preparation was CD4(187)-C4bp(SCR4), partially purified from conditioned culture fluid from a transformed CHO cell line on a FAST S® column. After infection at 37°C for 30 minutes, we added 15 µl aliquots in triplicate to 200 µl of RPMI-20% FCS in microtiter plates. We incubated the plates at 37°C in 5% CO_2 and examined them 4 to 8 days later for syncytia formation, a signal for active infection.

Our results indicated that, on a molar basis, multimeric CD4(187)-C4bp(SCR4) blocked HIV-1 infection at a concentration 100 times less than the concentration of recombinant soluble CD4 necessary to block HIV-1 infection. For example, 300 pM CD4(187)-C4bp(SCR4) (both partially purified and from cell culture medium) completely blocked syncytia formation. About 30 nM recombinant soluble CD4 was necessary to obtain the same result. There was a falloff in protection against HIV-1 infection of C8166 cells at about 100 pM for CD4(187)-C4bp(SCR4) and about 10 nM for recombinant soluble CD4.

C. CD4(187)-C4bp(SCR4) Blocks Splicing of HIV-1 mRNA In Vitro

We next carried out tests which demonstrated that CD4(187)-C4bp(SCR4) inhibits HIV-1 infection of cells. The assay we used measured the quantity of spliced HIV-1 mRNA produced when C8166 cells, HIV-1 and multimeric CD4-C4bp fusion proteins of this invention are incubated together.

Recombinant HIV-1 was obtained by transfecting colon carcinoma cell line SW480 with $CaPO_4$-precipitated pNL4-3. (Both the cell line and the plasmid are available from the AIDS Research and Reference Reagent Program, NIH, Bethesda, Maryland.) We incubated 10^7 C8166 cells with 10^3 $TCID_{50}$ recombinant HIV-1, alone or with serial dilutions of CD4(187)-C4bp(SCR4). After 48 hours, we determined the amount of spliced HIV-1 mRNA in total cellular RNA by nuclease S1 protection analysis.

We synthesized a 180 nucleotide single stranded DNA fragment probe with AMPLIGASE® (Epicenter Technologies), labelled at the 5' end with ^{32}P. The probe spanned the splice acceptor of all known spliced

HIV-1 mRNA molecules [B. Felber et al., "Feedback Regulation of Human Immunodeficiency Virus Type 1: Expression by the Rev Protein", J. Virol., 64, pp. 3734-41 (1990)]. We added this 180 nucleotide probe (10^3 cpm) to 10 μg of total cellular RNA in 10 μl of a buffer containing 80% formamide, 0.4 M NaCl, 40 mM PIPES, pH 6.4, 1 mM EDTA. We carried out the hybridization overnight at 48°C. After allowing hybridization, we treated the mixture with S1 nuclease. S1 nuclease completely digested any unhybridized probe and partially digested any hybridized probe, yielding a 102 nucleotide protected DNA fragment. The full length 180 nucleotide DNA fragment was also protected from digestion by hybridization to unspliced, genomic HIV-1 RNA, produced through replication of the virus in the cells. We found that the amount of both the 102 nucleotide DNA fragment and the 180 nucleotide fragment decreased with increasing amounts of C4bp fusion protein added as blocker. These results suggested that the CD4-C4bp fusion protein blocks HIV-1 entry into cells in concentrations as low as 1-10 ng/ml.

EXAMPLE VII -- PRODUCTION OF CD4(187)-C4bp(SCR8), CD4(187)-C4bp(SCR5), CD4(187)-C4bp(SCR3) AND CD4(187)-C4bp(SCR1)

We performed several ELISA assays to test for the production of CD4(187)-C4bp(SCR8), CD4(187)-C4bp(SCR5), CD4(187)-C4bp(SCR3) and CD4(187)-C4bp(SCR1). Our results indicated that COS-7 cells transformed with pJOD.sCD4.Y187.SCR8, pJOD.sCD4.Y187.SCR5, pJOD.sCD4.Y187.SCR3 and pJOD.sCD4.Y187.SCR1 all produced multimeric CD4-C4bp fusion proteins.

We performed four ELISA assays precisely as described previously (ELISAs 1-4) in the examples above, except that we used conditioned medium from

COS-7 cells transformed with pJOD.sCD4.Y187.SCR8. ELISAs 1-3 showed strong positive results. ELISA 4 (plates coated with gp120, 6C6 used as detection antibody) gave a weak positive result.

We performed six assays on conditioned medium from COS-7 cells transformed with three different plasmid isolates encoding CD4(187)-C4bp(SCR1). The isolates were designated, respectively, SCR1.1, SCR1.2 and SCR1.3. The first four assays were performed as described for ELISAs 1-4, above. We carried out the fifth assay (ELISA 5) in the same way as ELISA 4, except that we used the antibody 5A8 as the detection antibody. Antibody 5A8 does not block CD4 binding to gp120 and it recognizes domain 2 of CD4 (see Figure 4). Another monoclonal antibody having such characteristics might also be useful in this assay. We performed the sixth assay (ELISA 6) as in ELISA 3, except that we used the anti-C4bp monoclonals 051-198 or 051-28 (Quidel, San Diego, California).

In ELISA 5 (plate coated with gp120, 5A8 used as detection antibody) all three isolates gave positive results. This indicates that the cells produced a protein comprising CD4(187).

In ELISA 1 (plate coated with anti-hC4bp, 6C6 monoclonal used as detection antibody) all isolates gave negative results. In ELISA 3 (plate coated with gp120, anti-C4bp used as detection antibody) SCR1.1 gave a negative result and SCR1.2 and SCR1.3 gave a borderline positive result. In ELISA 6 (plate coated with gp120, 051-198 or 051-28 used as detection antibody) SCR1.2 and SCR1.3 gave positive results, but SCR1.1 gave negative results. This indicates that the cells produced molecules having both C4bp(SCR1) and the gp120 binding site of CD4. The reason that C4bp.SCR1 was not recognized in ELISA 1 and ELISA 3 is probably

due to the nature of the polyclonal antiserum and antibodies we used. These polyclonals seem to recognize preferentially epitopes on the mature form of human C4bp.

In ELISA 2 (plate coated with 6C6, 6C6 used as detection antibody) isolate SCR1.1 gave a negative result and isolates SCR1.2 and SCR1.3 gave positive results. In ELISA 4 (plate coated with gp120, 6C6 used as detection antibody) isolate SCR1.1 gave a negative result and isolates SCR1.2 and SCR1.3 gave positive results. These results are consistent with our belief that CD4(187)-C4bp(SCR1) can assemble into a multimer.

We performed another ELISA, ELISA 7, (plates coated with 1D7, 5A8 used as detection antibody) on conditioned medium from COS-7 cells transformed with pJOD.sCD4.Y187.SCR5.1, pJOD.sCD4.Y187.SCR5.2, pJOD.CD4.Y187.SCR3.2 and pJOD.sCD4.Y187.SCR3.3. All assays gave a strong positive result.

EXAMPLE VIII -- CLONING OF DNA ENCODING HBeAg-C4bp FUSION POLYPEPTIDES

We have constructed several plasmids characterized by DNA sequences encoding hepatitis B virus e antigen-C4bp ("HBeAg-C4bp") fusion polypeptides. We constructed these plasmids in two steps. First, we introduced a unique XbaI site into pJOD.C4bp.2, an isolate of pJOD.C4bp (Example I) between the DNA sequences encoding the C4bp signal sequence and the amino terminus of SCR8 (amino acid +1 of Figure 1). This created a site into which we could insert DNA sequences encoding HBeAg epitopes. Alternatively, one could employ this site to insert DNA sequences encoding other epitopes. In the second step, PCR fragments encoding various HBeAg sequences were inserted into the XbaI site to create polypeptides in which HBeAg sequences were sandwiched between the DNA

sequence encoding the signal sequence of C4bp and the DNA sequence encoding SCR8 of C4bp.

More specifically, in the first step, we introduced a unique XbaI site into pJOD.C4bp.2 via gapped mutagenesis as follows. We linearized a first sample of the plasmid with ApaI. Next, we cleaved a second sample of the plasmid with XhoI and SpeI. We then denatured the samples and allowed single stranded DNA from each to hybridize, creating a gap. We annealed a mutagenic oligomer into the gap. The oligomer had the sequence:

GAGGACCACAATTCTAGACCAAGAACAGCA.

We repaired the resulting plasmid with Klenow enzyme and dXPT, electroporated it into HB101 and isolated the plasmid, pJOD.C4bp.XbaI, which is characterized by a unique XbaI site and has, the sequence GGTCTAGAAT at the signal junction, with the GGT encoding the last amino acid of the C4bp signal sequence separated from the AAT encoding the first amino acid of SCR8 by the unique XbaI site.

We tested the plasmid as follows. We linearized pJOD.C4bp.XbaI with XbaI and blunt ended the fragment with mung bean nuclease. Then we religated the plasmid (without inserting any DNA) to create pJOD.C4bp.XbaI.O.3. The DNA sequence across the signal cleavage site thus became GGT AAT, encoding Gly(-1)Asn(1), as in authentic C4bp. When electroporated into COS cells, this plasmid expressed C4bp as efficiently as JOD.C4bp, as measured in a C4bp ELISA (ELISA 8: plates coated with rabbit anti-hC4bp, sheep anti-hC4bp used as detection antibody).

In the second step, we linearized pJOD.C4bp.XbaI with XbaI and blunt ended it with mung bean nuclease. We ligated PCR products encoding HBeAg(2-148), HBeAg(2-138), HBeAg(2-100) and HBeAg(2-

89), respectively, into the resultant vector. We used plasmid 8.1.5 as a template for the PCR fragments. Plasmid 8.1.5 is characterized by a DNA sequence encoding HBeAg. It is also referred to as pHBV139A [Pasek et al., supra]. (Plasmid 8.1.5 was a gift of Professor Kenneth Murray, University of Edinburgh, Scotland). Alternatively, one may use plasmid pAMG, ATCC 45020, as a PCR template. Plasmid pAMG is characterized by a DNA sequence derived from HBV subculture ADW and encodes HBeAg. We digested plasmid 8.1.5 with StyI, producing a fragment containing the entire PCR target. We performed PCR on the fragment using the following primers. We used the same 5' sense primer for all four constructs:

5' GACATTGACCCTTATAAAGAATTT.

The 3' anti-sense primers were:

5' AACAACAGTAGTCTCCGGAAGCGT [HBeAg(2-148)];

5' AGGGGCATTTGGTGGTCTATAAGC [HBeAg(2-138)];

5' TAATTGTCTGAACTTTAGGCCCAC [HBeAg(2-100)]; and

5' GACATAACTGACTACTAGGTCCCT [HBeAg(2-89)].

We ligated these PCR fragments with XbaI-digested, mung bean nuclease-treated pJOD.C4bp.XbaI. This ligation produced the following plasmids:

pJOD.HBeAg(2-148).C4bp.SCR8;

pJOD.HBeAg(2-138).C4bp.SCR8;

pJOD.HBeAg(2-100).C4bp.SCR8; and

pJOD.HBeAg(2-89).C4bp.SCR8, respectively.

These plasmids contained DNA sequences encoding the following fusion polypeptides: HBeAg(2-148)-C4bp(SCR8); HBeAg(2-138)-C4bp(SCR8); HBeAg(2-100)-C4bp(SCR8) and HBeAg(2-89)-C4bp(SCR8), respectively.

These constructs may be altered by replacing the C4bp signal sequence with the hepatitis B virus precore signal sequence to insure proper processing of the primary translation product.

Microorganisms and recombinant DNA molecules according to this invention are exemplified by cultures deposited in the In Vitro International, Inc. culture collection, in Linthicum, Maryland, USA on January 24, 1990, and identified as:

SCR1.1:	pJOD.sCD4.Y187.SCR1.1	IVI-10221
SCR1.2:	pJOD.sCD4.Y187.SCR1.2	IVI-10222
SCR1.3:	pJOD.sCD4.Y187.SCR1.3	IVI-10223
SCR8.2:	pJOD.sCD4.Y187.SCR8.2	IVI-10224
SCR8.3:	pJOD.sCD4.Y187.SCR8.3	IVI-10225
SCR4.2:	pJOD.sCD4.Y187.SCR4.2	IVI-10226
SCR4.3:	pJOD.sCD4.Y187.SCR4.3	IVI-10227
187.SnaB1:	pJOD.sCD4.Y187.SnaB1	IVI-10228
p170.2:	p170.2	IVI-10229
C4bp.3:	pJOD.C4bp.3	IVI-10230

We also deposited the following culture with In Vitro International on January 26, 1991:

Monoclonal Antibody 1D7-G11		IVI-10269

We also deposited the following cultures with In Vitro International on January 28, 1991:

SCR3.2:	pJOD.sCD4.Y187.SCR3.2	IVI-10270
SCR3.3:	pJOD.sCD4.Y187.SCR3.3	IVI-10271
SCR5.1:	pJOD.sCD4.Y187.SCR5.1	IVI-10272
SCR5.2:	pJOD.sCD4.Y187.SCR5.2	IVI-10273

While we have hereinbefore described a number of embodiments of this invention, it is apparent that our basic embodiments can be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the scope of this invention includes all alternative embodiments and variations which are defined in the foregoing specification and by the claims appended hereto; and the invention is not to be limited by the specific embodiments which have been presented herein by way of example.

CLAIMS

We claim:

1. A recombinant DNA molecule comprising a DNA sequence encoding a C4bp fusion polypeptide.

2. The recombinant DNA molecule according to claim 1, wherein the DNA sequence comprises a DNA sequence encoding a polypeptide moiety fused to the 5' end of a DNA sequence encoding a C4bp monomer.

3. The recombinant DNA molecule according to claim 2, wherein the C4bp monomer comprises at most eight SCRs.

4. The recombinant DNA molecule according to claim 3, wherein the C4bp monomer is selected from the group consisting of a C4bp monomer having eight SCRs and comprising amino acids +1 to +549 of Figure 1, a C4bp monomer having five SCRs and comprising amino acids +188 to +549 of Figure 1, a C4bp monomer having four SCRs and comprising amino acids +248 to +549 of Figure 1, a C4bp monomer having three SCRs and comprising amino acids +314 to +549 of Figure 1, and a C4bp monomer having one SCR and comprising amino acids +433 to +549 of Figure 1.

5. The recombinant DNA molecule according to claim 2, wherein the polypeptide moiety is selected from the group consisting of viral receptors, cell receptors, cell ligands, bacterial immunogens, parasitic immunogens, viral immunogens, immunoglobulins or fragments thereof that bind to target molecules, enzymes, enzyme inhibitors, enzyme substrates,

cytokines, growth factors, colony stimulating factors, hormones and toxins.

6. The recombinant DNA molecule according to claim 5, wherein the polypeptide moiety is a soluble CD4 protein.

7. The recombinant DNA molecule according to claim 6, wherein the soluble CD4 protein is selected from the group consisting of CD4(111), CD4(181), CD4(183), CD4(187) and CD4(375).

8. The recombinant DNA molecule according to claim 5, wherein the polypeptide moiety comprises a viral polypeptide displaying hepatitis B virus e antigenicity.

9. The recombinant DNA molecule according to claim 8, wherein the viral polypeptide is selected from the group consisting of HBeAg(2-89), HBeAg(2-100), HBeAg(2-138) and HBeAg(2-148).

10. The recombinant DNA molecule according to claim 5, wherein the polypeptide moiety is a cell receptor or a cell ligand selected from the group consisting of ICAM1, ELAM1, VCAM1 or VCAM1b and LFA3.

11. The recombinant DNA molecule according to claim 5, wherein the polypeptide moiety is selected from the group consisting of hirudin, C-terminal hirudin peptides and hirulogs.

12. The recombinant DNA molecule according to claim 1 or 2, wherein the DNA sequence encoding a

C4bp fusion polypeptide is operatively linked to an expression control sequence.

13. The recombinant DNA molecule according to claim 12, wherein the expression control sequence is selected from the group consisting of the early and late promoters of SV40 or adenovirus, the lac system, the trp system, the TAC or TRC system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, the promoters of the yeast α-mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof.

14. The recombinant DNA molecule according to claim 13, said molecule being selected from the group consisting of pJOD.sCD4.Y187.SCR1.1, pJOD.sCD4.Y187.SCR1.2, pJOD.sCD4.Y187.SCR1.3, pJOD.sCD4.Y187.SCR3.2, pJOD.sCD4.Y187.SCR3.3, pJOD.sCD4.Y187.SCR4.2, pJOD.sCD4.Y187.SCR4.3, pJOD.sCD4.Y187.SCR5.1, pJOD.sCD4.Y187.SCR5.2, pJOD.sCD4.Y187.SCR8.2 and pJOD.sCD4.Y187.SCR8.3.

15. The recombinant DNA molecule according to claim 13, selected from the group consisting of pJOD.HBeAg(2-89).C4bp.SCR8, pJOD.HBeAg(2-100).C4bp.SCR8, pJOD.HBeAg(2-138).C4bp.SCR8 and pJOD.HBeAg(2-148).C4bp.SCR8.

16. A unicellular host transformed with a recombinant DNA molecule according to claim 12.

17. The unicellular host according to claim 16, said molecule being selected from the group consisting of E.coli, Pseudomonas, Bacillus, Streptomyces, fungi, yeasts, CHO cells, mouse cells, African green monkey cells, COS-1 cells, COS-7 cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, insect cells, and human cells and plant cells in tissue culture.

18. The unicellular host according to claim 17, said host being a COS-7 cell or a CHO cell.

19. The unicellular host according to claim 17, said host being transformed with pJOD.sCD4.Y187.SCR4.

20. A C4bp fusion polypeptide comprising a functional moiety and a C4bp monomer.

21. The C4bp fusion polypeptide according to claim 20, wherein the C4bp monomer is selected from the group consisting of a C4bp monomer having eight SCRs and comprising amino acids +1 to +549 of Figure 1, a C4bp monomer having five SCRs and comprising amino acids +188 to +549 of Figure 1, a C4bp monomer having four SCRs and comprising amino acids +248 to +549 of Figure 1, a C4bp monomer having three SCRs and comprising amino acids +314 to +549 of Figure 1, and a C4bp monomer having one SCR and comprising amino acids +433 to +549 of Figure 1.

22. The C4bp fusion polypeptide according to claim 20, wherein the functional moiety is selected from the group consisting of viral receptors, cell receptors, cell ligands, bacterial immunogens,

parasitic immunogens, viral immunogens, immunoglobulins or fragments thereof that bind to target molecules, enzymes, enzyme inhibitors, enzyme substrates, cytokines, growth factors, colony stimulating factors, hormones and toxins.

23. The C4bp fusion polypeptide according to claim 22, wherein the functional moiety is a soluble CD4 protein.

24. The C4bp fusion polypeptide according to claim 23, wherein the soluble CD4 protein is selected from the group consisting of CD4(111), CD4(181), CD4(183), CD4(187)-C4bp(SCR 5), CD4(187)-C4bp(SCR3) and CD4(375).

25. The C4bp fusion polypeptide according to claim 23, said polypeptide being selected from the group consisting of CD4(187)-C4bp(SCR8), CD4(187)-C4bp(SCR5), CD4(187)-C4bp(SCR4), CD4(187)-C4bp(SCR3) and CD4(187)-C4bp(SCR1).

26. The C4bp fusion polypeptide according to claim 22, wherein the functional moiety is a viral polypeptide displaying hepatitis B virus e antigenicity.

27. The C4bp fusion polypeptide according to claim 26, said polypeptide selected from the group consisting of HBeAg(2-89)-C4bp(SCR8), HBeAg(2-100)-C4bp(SCR8), HBeAg(2-138)-C4bp(SCR8) and HBeAg(2-148)-C4bp(SCR8).

28. The C4bp fusion polypeptide according to claim 22, wherein said functional moiety is a cell

receptor or a cell ligand selected from the group consisting of ICAM1, ELAM1, VCAM1 or VCAM1b and LFA3.

29. The C4bp fusion polypeptide according to claim 22, wherein the functional moiety is selected from the group consisting of hirudin, C-terminus hirudin polypeptides and hirulogs.

30. The C4bp fusion polypeptide according to claim 20, wherein the C-terminus of the polypeptide moiety is fused to the N-terminus of the C4bp monomer.

31. The C4bp fusion polypeptide according to claim 20, said functional moiety being selected from the group consisting of toxins, anti-retroviral agents, enzyme substrates and enzyme inhibitors.

32. The C4bp fusion polypeptide according to claim 31, wherein the functional moiety is AZT.

33. The C4bp fusion polypeptide according to claim 20, wherein said functional moiety comprises a reporter group selected from the group consisting of enzymes, radionuclides, fluorescent markers and chemiluminescent markers.

34. A multimeric C4bp fusion protein.

35. The fusion protein according to claim 34, said protein being a multimeric CD4-C4bp fusion protein.

36. The fusion protein according to claim 35, said protein being CD4(187)-C4bp(SCR4) fusion protein.

37. The fusion protein according to claim 34, said protein being a multimeric HBeAg-C4bp fusion protein.

38. The fusion protein according to claim 34, said protein being selected from the group consisting of ELAM1-C4bp fusion proteins, VCAM1-C4bp fusion proteins, VCAM1b-C4bp fusion proteins and ICAM1-C4bp fusion proteins.

39. The fusion protein according to claim 34, said protein being selected from the group consisting of hirudin-C4bp fusion proteins, C-terminus hirudin polypeptide-C4bp fusion proteins and hirulog-C4bp fusion proteins.

40. A method for producing a C4bp fusion polypeptide comprising the step of transforming a unicellular host with a recombinant DNA molecule comprising a DNA sequence encoding a C4bp fusion polypeptide operatively linked to an expression control sequence.

41. A hetero-multimeric C4bp fusion protein.

42. The hetero-multimeric C4bp fusion protein according to claim 41, said fusion protein comprising a first functional moiety selected from the group consisting of viral receptors, cell receptors and cell ligands, and a second functional moiety selected from the group consisting of toxins and anti-retroviral agents.

43. The hetero-multimeric C4bp fusion protein according to claim 41, said fusion protein comprising a recognition molecule and a reporter group.

44. The hetero-multimeric C4bp fusion protein according to claim 41, wherein the first functional moiety is soluble CD4 and the second functional moiety is AZT.

45. The hetero-multimeric C4bp fusion protein according to claim 41, said fusion protein comprising at least two different immunogens.

46. A method for producing a multimeric C4bp fusion protein comprising the step of transforming a unicellular host with a recombinant DNA molecule of claim 12.

47. A method for treating a patient having AIDS, ARC, HIV infection or antibodies to HIV comprising the step of administering to the patient a therapeutically effective amount of a multimeric CD4-C4bp fusion protein of claim 35 or a hetero-multimeric CD4-C4bp fusion protein of claim 42.

48. The method according to claim 47, wherein the fusion protein comprises CD4(187)-C4bp(SCR4).

49. The method according to claim 47, wherein the hetero-multimeric CD4-C4bp fusion protein is a fusion protein according to claim 44.

50. A method for identifying the presence of a target molecule in a sample comprising the step of

contacting the sample with a hetero-multimeric C4bp fusion protein according to claim 43.

51. A method for identifying the presence of a target molecule in vivo comprising the step of administering to a patient an effective amount of a hetero-multerimic C4bp fusion protein according to claim 41.

52. A method for treating human disease comprising the step of infecting human somatic cells with a retrovirus comprising a DNA sequence encoding a C4bp fusion polypeptide.

53. The method according to claim 52, wherein said DNA sequence encodes a CD4-C4bp fusion polypeptide.

54. A recombinant human C4 binding protein.

55. A method for producing recombinant C4 binding protein comprising the step of transforming a unicellular host with a recombinant DNA molecule comprising an expression control sequence operatively linked to a DNA sequence comprising the DNA sequence of Figure 1 from nucleotide 4 to nucleotide 1743.

56. A recombinant DNA molecule comprising a DNA sequence encoding a non-human C4bp fusion polypeptide.

57. The recombinant DNA molecule according to claim 56, wherein the non-human is a mouse or a guinea pig.

58. A unicellular host transformed with a recombinant DNA molecule according to claim 56, wherein the DNA sequence encodes a non-human C4bp fusion polypeptide.

59. The unicellular host according to claim 58, wherein the non-human is a mouse or a guinea pig.

60. A non-human C4bp fusion polypeptide comprising a functional moiety and a non-human C4bp monomer.

61. The C4bp fusion polypeptide according to claim 60, wherein the non-human is a mouse or a guinea pig.

62. A multimeric non-human C4bp fusion protein.

63. The multimeric non-human C4bp fusion protein according to claim 62, wherein the non-human is a mouse or a guinea pig.

64. A hetero-multimeric non-human C4bp fusion protein.

65. The hetero-multimeric non-human C4bp fusion protein according to claim 64, wherein the non-human is a mouse or a guinea pig.

66. A method for producing a non-human C4 binding protein comprising the step of transforming a unicellular organism with a recombinant DNA molecule of claim 56.

67. The method according to claim 66, wherein the non-human is a guinea pig or a mouse.

FIG. 1A

```
   1  ATGGCAGCCTGGCCCTTCTCCAGGCTGTGGAAAGTCTCTGATCCAATTCTCTTCCAAATG   60
- 32  MetAlaAlaTrpProPheSerArgLeuTrpLysValSerAspProIleLeuPheGlnMet  - 13

  61  ACCTTGATCGCTGCTCTGTTGCCTGCTGTTCTTGGCAATTGTGGTCCTCCACCCACTTTA  120
- 12  ThrLeuIleAlaAlaLeuLeuProAlaValLeuGlyAsnCysGlyProProProThrLeu    8

 121  TCATTTGCTGCCCCGATGGATATTACGTTGACTGAGACACGCTTCAAAACTGGAACTACT  180
   9  SerPheAlaAlaProMetAspIleThrLeuThrGluThrArgPheLysThrGlyThrThr   28

 181  CTGAAATACACCTGCCTCCCTGGCTACGTCAGATCCCATTCAACTCAGACGCTTACCTGT  240
  29  LeuLysTyrThrCysLeuProGlyTyrValArgSerHisSerThrGlnThrLeuThrCys   48

 241  AATTCTGATGGCGAATGGGTGTATAACACCTTCTGTATCTACAAACGATGCAGACACCCA  300
  49  AsnSerAspGlyGluTrpValTyrAsnThrPheCysIleTyrLysArgCysArgHisPro   68
```

FIG. 1B

```
301  GGAGAGTTACGTAATGGGCAAGTAGAGATTAAGACAGATTTATCTTTTGGATCACAAATA  360
 69  GlyGluLeuArgAsnGlyGlnValGluIleLysThrAspLeuSerPheGlySerGlnIle   88

361  GAATTCAGCTGTTCAGAAGGATTTTTCTTAATTGGCTCAACCACTAGTCGTTGTGAAGTC  420
 89  GluPheSerCysSerGluGlyPhePheLeuIleGlySerThrThrSerArgCysGluVal  108

421  CAAGATAGAGGAGTTGGCTGGAGTCATCCTCTCCCACAATGTGAAATTGTCAAGTGTAAG  480
109  GlnAspArgGlyValGlyTrpSerHisProLeuProGlnCysGluIleValLysCysLys  128

481  CCTCCTCCAGACATCAGGAATGGAAGGCACAGCGGTGAAGAAAATTTCTACGCATACGGC  540
129  ProProProAspIleArgAsnGlyArgHisSerGlyGluGluAsnPheTyrAlaTyrGly  148

541  TTTTCTGTCACCTACAGCTGTGACCCCCGCTTCTCACTCTTGGGCCATGCCTCCATTTCT  600
149  PheSerValThrTyrSerCysAspProArgPheSerLeuLeuGlyHisAlaSerIleSer  168
```

FIG. IC

```
601  TGCACTGTGGAGAATGAAACAATAGGCGTTTGGAGACCAAGCCCTCCTACCTGTGAAAAA  660
169  CysThrValGluAsnGluThrIleGlyValTrpArgProSerProProThrCysGluLys  188

661  ATCACCTGTCGCAAGCCAGATGTTTCACATGGGGAAATGGTCTCTGGATTTGGACCCATC  720
189  IleThrCysArgLysProAspValSerHisGlyGluMetValSerGlyPheGlyProIle  208

721  TATAATTACAAAGACACTATTGTGTTTAAGTGCCAAAAAGGTTTTGTTCTCAGAGGCAGC  780
209  TyrAsnTyrLysAspThrIleValPheLysCysGlnLysGlyPheValLeuArgGlySer  228

781  AGTGTAATTCATTGTGATGCTGATAGCAAATGGAATCCTTCTCCTCCTGCTTGTGAGCCC  840
229  SerValIleHisCysAspAlaAspSerLysTrpAsnProSerProProAlaCysGluPro  248

841  AATAGTTGTATTAATTTACCAGACATTCCACATGCTTCCTGGGAAACATATCCTAGGCCG  900
249  AsnSerCysIleAsnLeuProAspIleProHisAlaSerTrpGluThrTyrProArgPro  268
```

FIG. ID

```
 901  ACAAAAGAGGATGTGTATGTTGTTGGGACTGTGTTAAGGTACCGCTGTCATCCTGGCTAC  960
 269  ThrLysGluAspValTyrValValGlyThrValLeuArgTyrArgCysHisProGlyTyr  288

 961  AAACCCACTACAGATGAGCCTACGACTGTGATTTGTCAGAAAAATTTGAGATGGACCCCA 1020
 289  LysProThrThrAspGluProThrThrValIleCysGlnLysAsnLeuArgTrpThrPro  308

1021  TACCAAGGATGTGAGGCGTTATGTTGCCCTGAACCAAAGCTAAATAATGGTGAAATCACT 1080
 309  TyrGlnGlyCysGluAlaLeuCysCysProGluProLysLeuAsnAsnGlyGluIleThr  328

1081  CAACACAGGAAAAGTCGTCCTGCCAATCACTGTGTTTATTTCTATGGAGATGAGATTTCA 1140
 329  GlnHisArgLysSerArgProAlaAsnHisCysValTyrPheTyrGlyAspGluIleSer  348

1141  TTTTCATGTCATGAGACCAGTAGGTTTTCAGCTATATGCCAAGGAGATGGCACGTGGAGT 1200
 349  PheSerCysHisGluThrSerArgPheSerAlaIleCysGlnGlyAspGlyThrTrpSer  368
```

FIG. IE

```
1201  CCCCGAACACCATCATGTGGAGACATTTGCAATTTTCCTCCTAAAATTGCCCATGGGCAT  1260
 369  ProArgThrProSerCysGlyAspIleCysAsnPheProProLysIleAlaHisGlyHis   388

1261  TATAAACAATCTAGTTCATACAGCTTTTTCAAAGAAGAGATTATATATGAATGTGATAAA  1320
 389  TyrLysGlnSerSerSerTyrSerPhePheLysGluGluIleIleTyrGluCysAspLys   408

1321  GGCTACATTCTGGTCGGACAGGCGAAACTCTCCTGCAGTTATTCACACTGGTCAGCTCCA  1380
 409  GlyTyrIleLeuValGlyGlnAlaLysLeuSerCysSerTyrSerHisTrpSerAlaPro   428

1381  GCCCCTCAATGTAAAGCTCTGTGTCGGAAACCAGAATTAGTGAATGGAAGGTTGTCTGTG  1440
 429  AlaProGlnCysLysAlaLeuCysArgLysProGluLeuValAsnGlyArgLeuSerVal   448

1441  GATAAGGATCAGTATGTTGAGCCTGAAAATGTCACCATCCAATGTGATTCTGGCTATGGT  1500
 449  AspLysAspGlnTyrValGluProGluAsnValThrIleGlnCysAspSerGlyTyrGly   468
```

FIG. IF

```
1501  GTGGTTGGTCCCCAAAGTATCACTTGCTCTGGGAACAGAACCTGGTACCCAGAGGTGCCC  1560
 469  ValValGlyProGlnSerIleThrCysSerGlyAsnArgThrTrpTyrProGluValPro   488

1561  AAGTGTGAGTGGGAGACCCCCGAAGGCTGTGAACAAGTGCTCACAGGCAAAAGACTCATG  1620
 489  LysCysGluTrpGluThrProGluGlyCysGluGlnValLeuThrGlyLysArgLeuMet   508

1621  CAGTGTCTCCCAAACCCAGAGGATGTGAAAATGGCCCTGGAGGTATATAAGCTGTCTCTG  1680
 509  GlnCysLeuProAsnProGluAspValLysMetAlaLeuGluValTyrLysLeuSerLeu   528

1681  GAAATTGAACAACTGGAACTACAGAGAGACAGCGCAAGACAATCCACTTTGGATAAAGAA  1740
 529  GluIleGluGlnLeuGluLeuGlnArgAspSerAlaArgGlnSerThrLeuAspLysGlu   548

1741  CTATAA  1746
 549  LeuEnd  550
```

FIG. 2

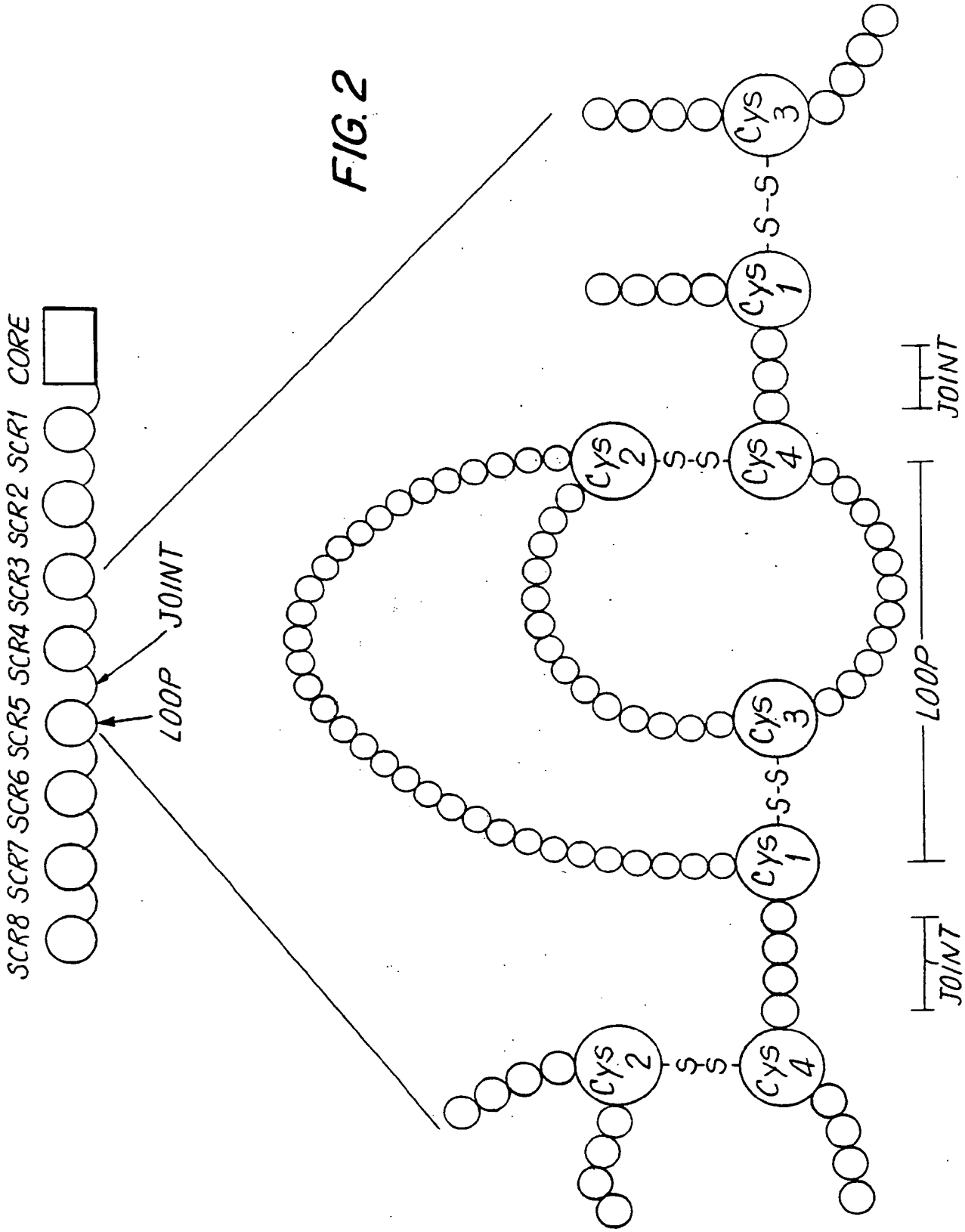

SCR8 SCR7 SCR6 SCR5 SCR4 SCR3 SCR2 SCR1 CORE
JOINT
LOOP
CYS 3
S-S
CYS 1
CYS 2
S-S
CYS 4
CYS 3
S-S
CYS 1
CYS 2
S-S
CYS 4
JOINT
LOOP
JOINT

FIG. 3A

```
   1  ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCA   60
 -25  MetAsnArgGlyValProPheArgHisLeuLeuLeuValLeuGlnLeuAlaLeuLeuPro   -6

  61  GCAGCCACTCAGGGAAAGAAAGTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACC  120
  -5  AlaAlaThrGlnGlyLysLysValValLeuGlyLysLysGlyAspThrValGluLeuThr   15

 121  TGTACAGCTTCCCAGAAGAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAG  180
  16  CysThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLys   35

 181  ATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATCCAAGCTGAATGATCGCGCT  240
  36  IleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSerLysLeuAsnAspArgAla   55

 241  GACTCAAGAAGAAGCTTGTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAG  300
  56  AspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLys   75
```

8/30

FIG. 3B

```
301  ATAGAAGACTCAGATACTTACATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTG  360
 76  IleGluAspSerAspThrTyrIleCysGluValGluAspGlnLysGluGluValGlnLeu   95

361  CTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACC  420
 96  LeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThr  115

421  CTGACCTTGGAGAGCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAGGGGT  480
116  LeuThrLeuGluSerProProGlySerSerProSerValGlnCysArgSerProArgGly  135

481  AAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGC  540
136  LysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeuGluLeuGlnAspSerGly  155

541  ACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAAATAGACATCGTG  600
156  ThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleVal  175
```

9/30

FIG. 3C

```
601  GTGCTAGCTTTCCAGAAGGCCTCCAGCATAGTCTACAAGAAAGAGGGGGAACAGGTGGAG  660
176  ValLeuAlaPheGlnLysAlaSerSerIleValTyrLysLysGluGlyGluGlnValGlu  195

661  TTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGG  720
196  PheSerPheProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrp  215

721  CAGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTCTGACCTGAAGAACAAGGAA  780
216  GlnAlaGluArgAlaSerSerSerLysSerTrpIleThrSerAspLeuLysAsnLysGlu  235

781  GTGTCTGTAAAACGGGTTACCCAGGACCCTAAGCTCCAGATGGGCAAGAAGCTCCCGCTC  840
236  ValSerValLysArgValThrGlnAspProLysLeuGlnMetGlyLysLysLeuProLeu  255

841  CACCTCACCCTGCCCCAGGCCTTGCCTCAGTATGCTGGCTCTGGAAACCTCACCCTGGCC  900
256  HisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAla  275
```

FIG. 3D

```
 901  CTTGAAGCGAAAACAGGAAAGTTGCATCAGGAAGTGAACCTGGTGGTGATGAGAGCCACT  960
 276  LeuGluAlaLysThrGlyLysLeuHisGlnGluValAsnLeuValValMetArgAlaThr  295

 961  CAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTG 1020
 296  GlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeu  315

1021  AGCTTGAAACTGGAGAACAAGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTG 1080
 316  SerLeuLysLeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpVal  335

1081  CTGAACCCTGAGGCGGGGATGTGGCAGTGTCTGCTGAGTGACTCGGGACAGGTCCTGCTG 1140
 336  LeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAspSerGlyGlnValLeuLeu  355
```

FIG. 3E

```
1141  GAATCCAACATCAAGGTTCTGCCCACATGGTCCACCCCGGTGCAGCCAATGGCCCTGATT  1200
 356  GluSerAsnIleLysValLeuProThrTrpSerThrProValGlnProMetAlaLeuIle   375

1201  GTGCTGGGGGGCGTCGCCGGCCTCCTGCTTTTCATTGGGCTAGGCATCTTCTTCTGTGTC  1260
 376  ValLeuGlyGlyValAlaGlyLeuLeuLeuPheIleGlyLeuGlyIlePhePheCysVal   395

1261  AGGTGCCGGCACCGAAGGCGCCAAGCAGAGCGGATGTCTCAGATCAAGAGACTCCTCAGT  1320
 396  ArgCysArgHisArgArgArgGlnAlaGluArgMetSerGlnIleLysArgLeuLeuSer   415

1321  GAGAAGAAGACCTGCCAGTGCCCTCACCGGTTTCAGAAGACATGTAGCCCCATTTGA     1377
 416  GluLysLysThrCysGlnCysProHisArgPheGlnLysThrCysSerProIleEnd      434
```

FIG. 4

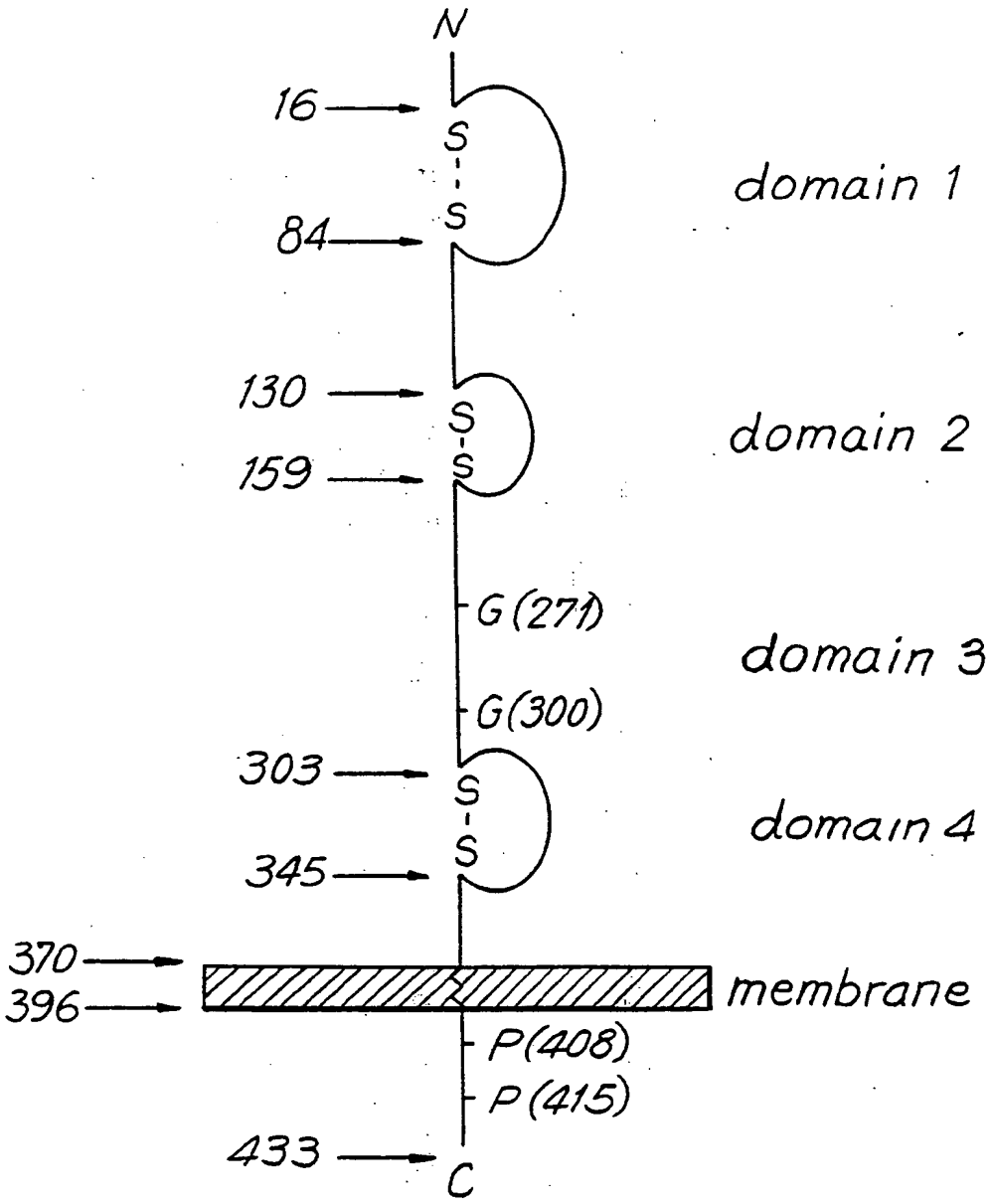

N
16
S
S
84
domain 1
130
S
S
159
domain 2
G(271)
G(300)
domain 3
303
S
S
345
domain 4
370
396
membrane
P(408)
P(415)
433
C

FIG. 5A

C4bp.1	ATGGCAGCCT GGCCCTTCTC CAGG
C4bp.2	TTATAGTTCT TTATCCAAAG TGGA
C4bp.3	CAAGACACCT TTTCCTCCTT CTTTTGAGAA
C4bp.4	CTCGAGGAAT TCCCCATGGC AGCCTGGCCC
C4bp.5	GATAAAGAAC TATAAGGGAC ATATGACGTG
C4bp.6	CATAAATTGG TCTGCTCGAG CCACCATGCA
C4bp.7	CTAGCTTTCC AGAAGGCCTC CAGCATAGTC TACGTAA

FIG. 5B

C4bp.8	GATCTTACGT AGACTATGCT GGAGGCCTTC TGGAAAG
C4bp.9	ATCCAAGCTG GTCGAATGGC AGCCTGGCCC
C4bp.10	GATAAAGAAC TATAATCGAC CGTGACCCCT
C4bp.11	GCATAGTCTA CGTAAGATCT TTGTGAAGGA
C4bp.12	TCCAGCATAG TCTACAATTG TGGTCCTCCA
C4bp.13	TCCAGCATAG TCTACCCCAA TAGTTGTATT
C4bp.14	TCCAGCATAG TCTACTTAGT GAATGGAAGG

FIG. 5C

C4bp.15	CTAGCTTTCC AGGTACCGTC GAC
C4bp.16	TCGACGGTAC CTGGAAAGCT AGC
C4bp.17	TCCAGCATAG TCTACAAAGC TCTGTGCCGG
C4bp.18	CTAGCTTTCC AGGTACCG
C4bp.19	TCGACGGTAC CTGGAAAG
C4bp.20	TTACTCACAC TTGGGCACCT CTGG

FIG. 5D

SCR.8	AATTGTGGTC CTCCACCCAC TTTA
SCR.4	CCCAATAGTT GTATTAATTT ACCA
SCR.1	AAAGCTCTGT GCCGGAAACC AGAA
312.20	GCGTTATGTT GCCCTGAACC AAAG
312.21	AAATCACCT GTCGCAAGCC AGAT
312.35	TCCAGCATAG TCTACGCGTT ATGTTGCCCT
312.36	TCCAGCATAG TCTACAAAAT CACCTGTCGC

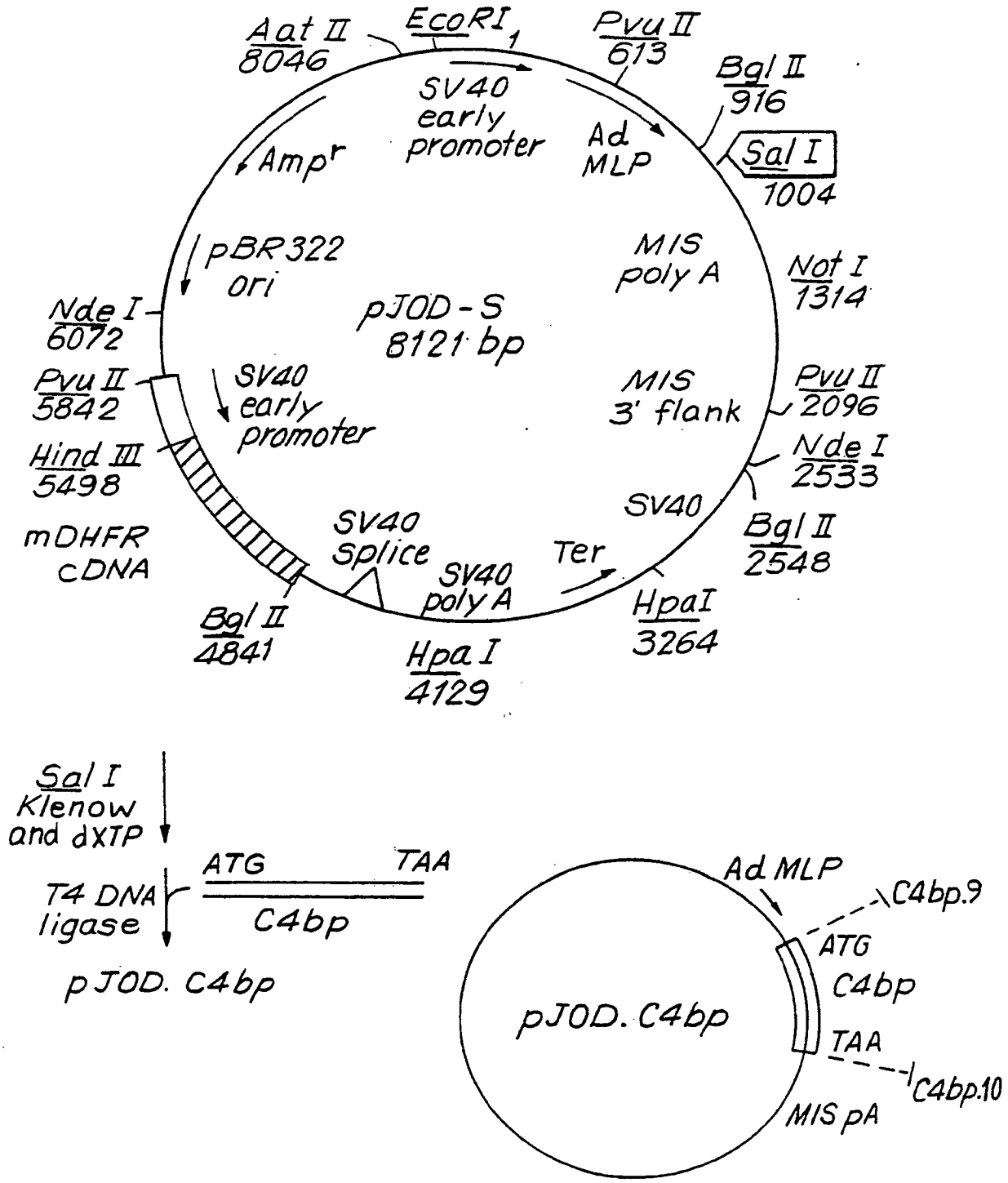

Aat II 8046
EcoRI 1
Pvu II 613
Bgl II 916
Sal I 1004
Not I 1314
Pvu II 2096
Nde I 2533
Bgl II 2548
HpaI 3264
Hpa I 4129
Bgl II 4841
Pvu II 5842
Hind III 5498
Nde I 6072
SV40 early promoter
Ampr
Ad MLP
MIS poly A
MIS 3' flank
SV40
Ter
SV40 Splice
SV40 poly A
pBR322 ori
pJOD-S 8121 bp
mDHFR cDNA
Sal I Klenow and dXTP
T4 DNA ligase
ATG
TAA
C4bp
pJOD. C4bp
Ad MLP
C4bp.9
ATG
C4bp
TAA
C4bp.10
MIS pA
pJOD. C4bp

FIG. 6A

FIG. 6B

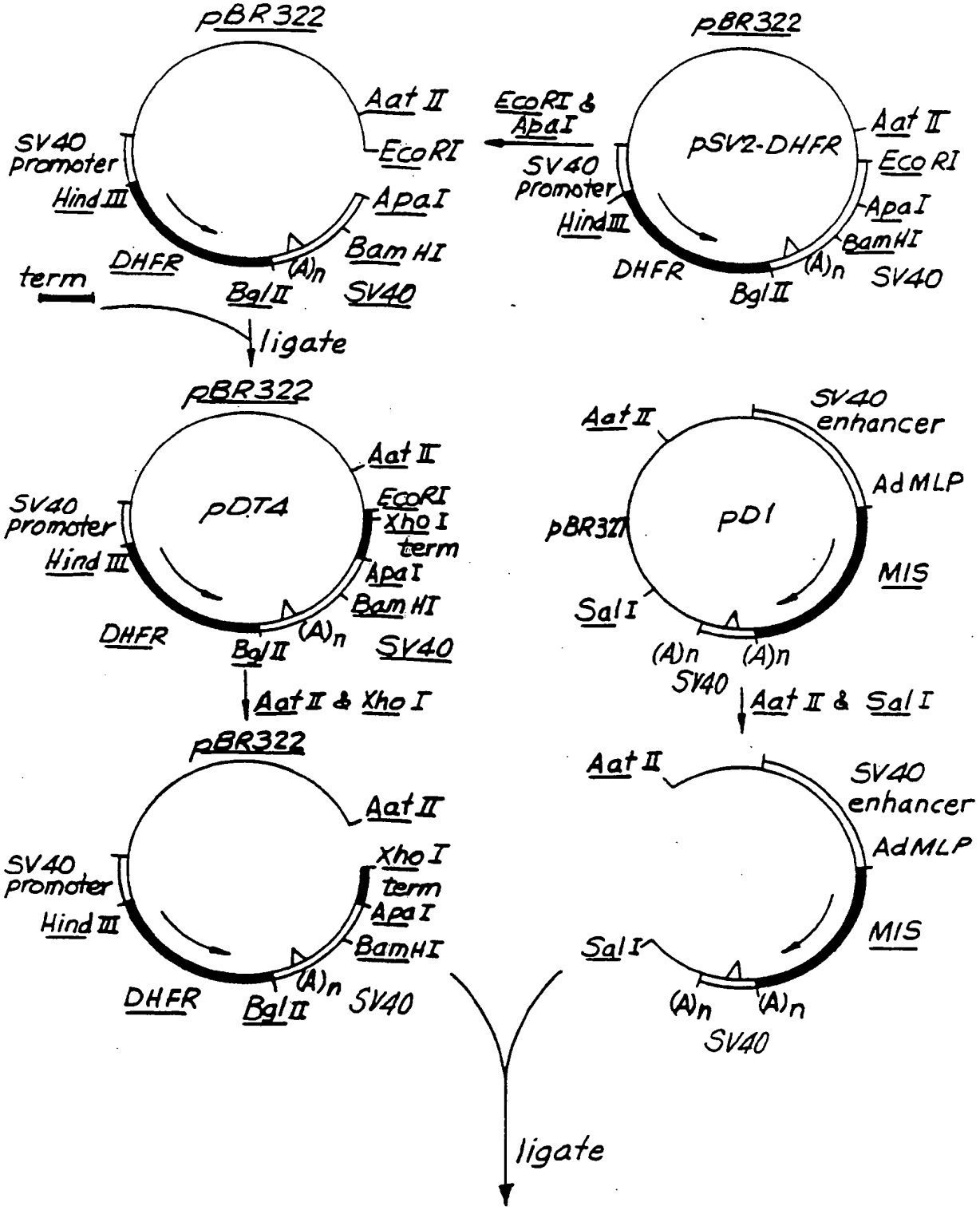


FIG. 6B (con't)

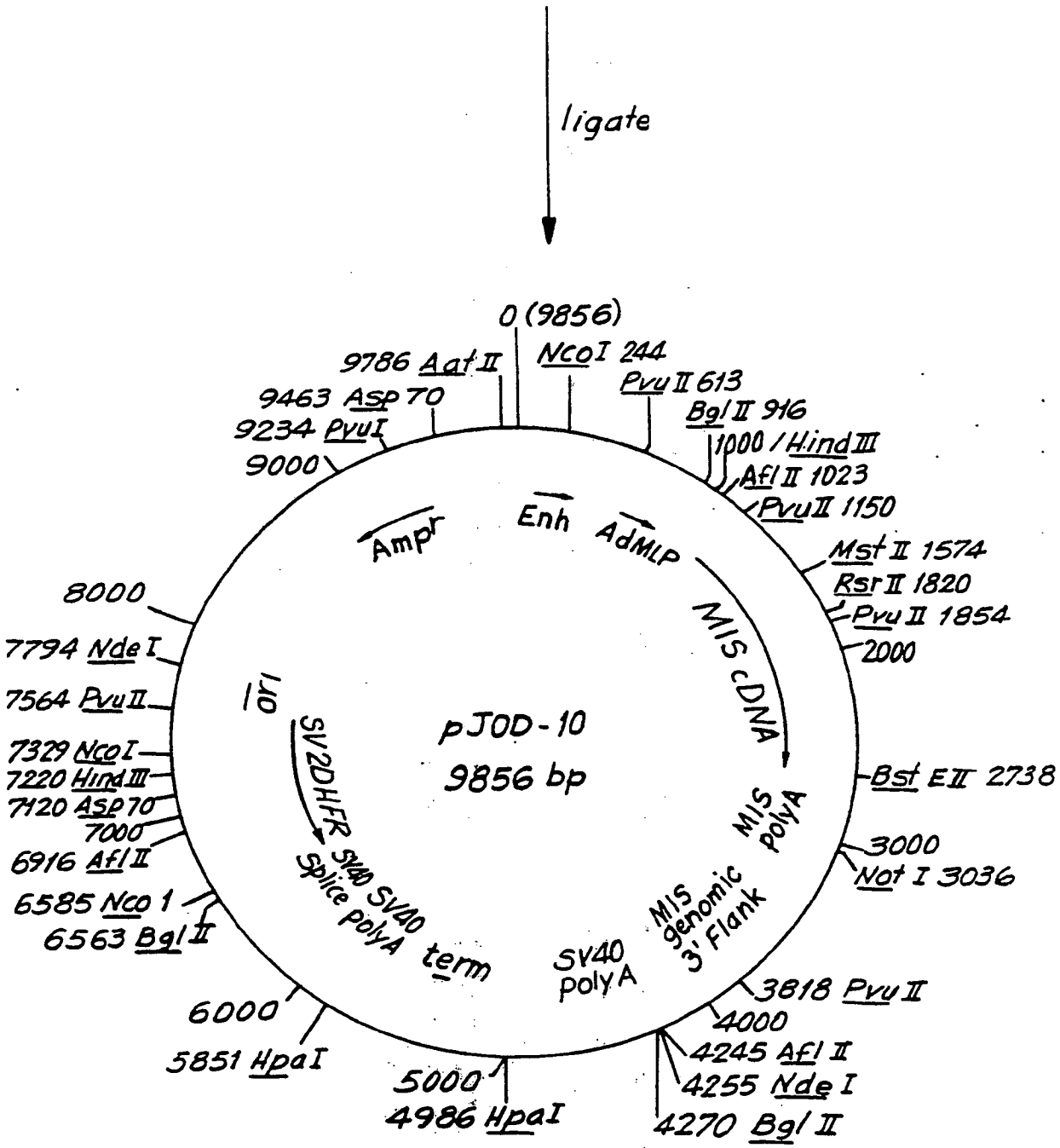


FIG. 6C

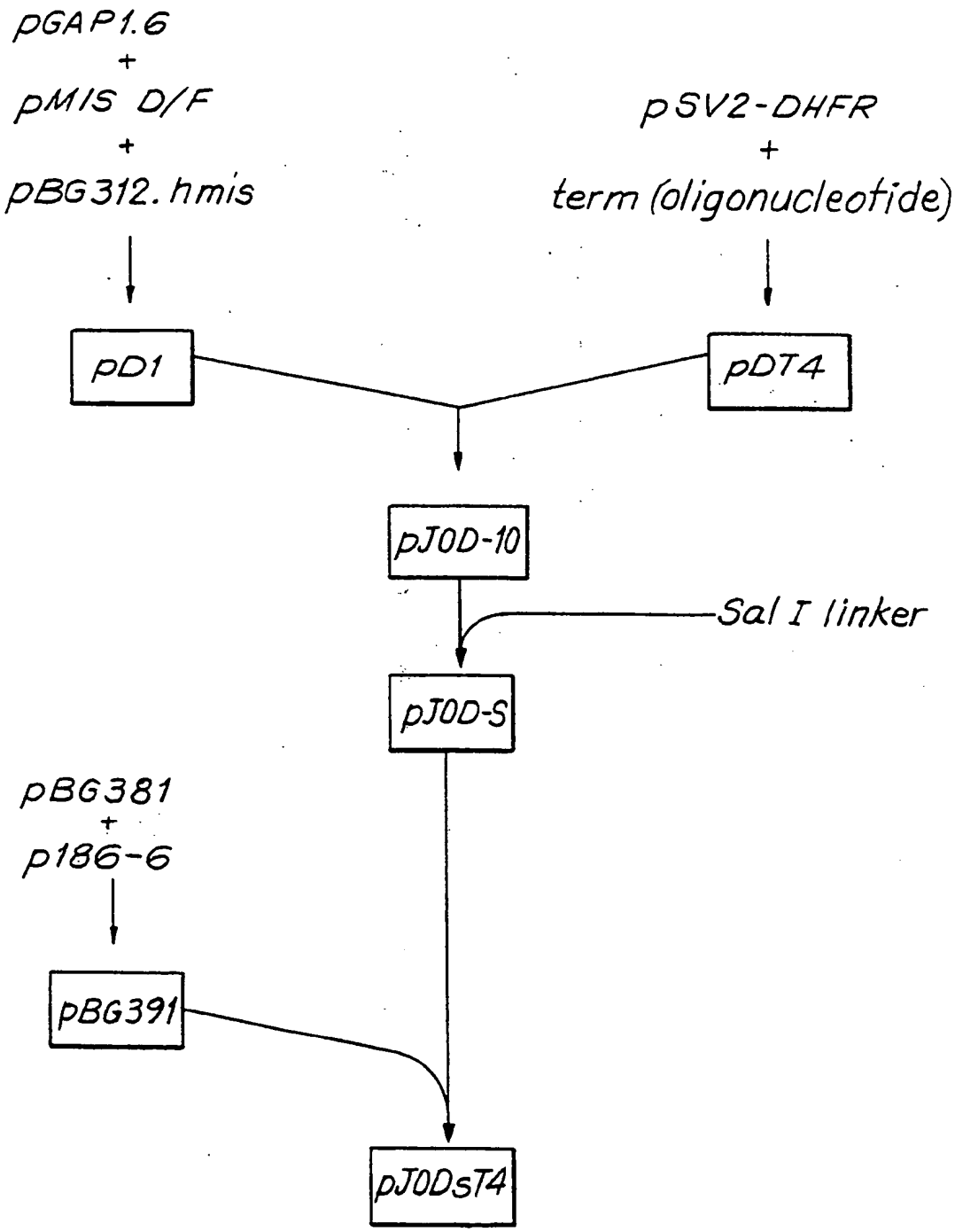
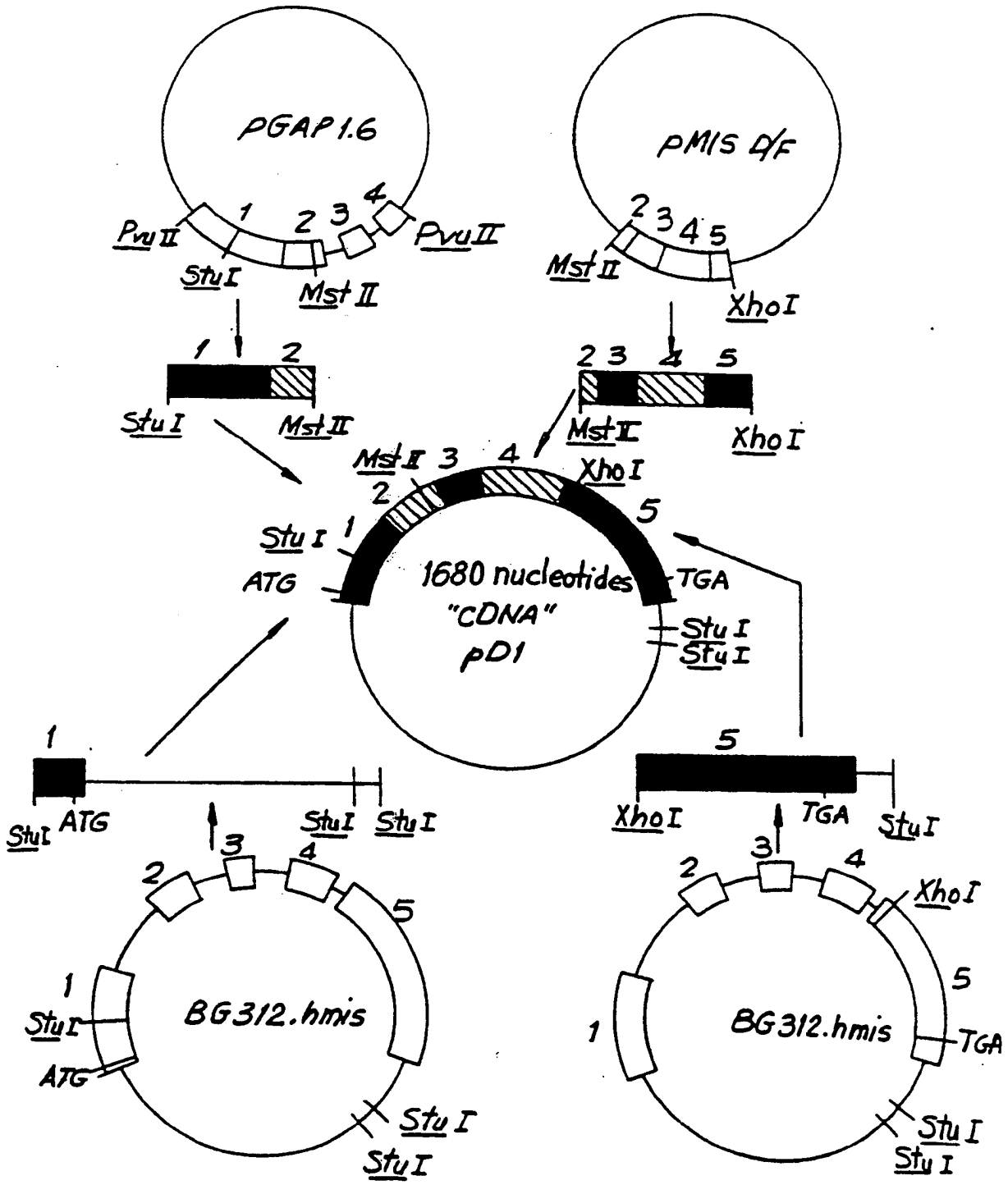


FIG. 6D


pGAP1.6
PvuII
PvuII
StuI
MstII
pMIS D/F
MstII
XhoI
1680 nucleotides
"cDNA"
pD1
ATG
TGA
StuI
StuI
BG312.hmis
BG312.hmis
XhoI
TGA

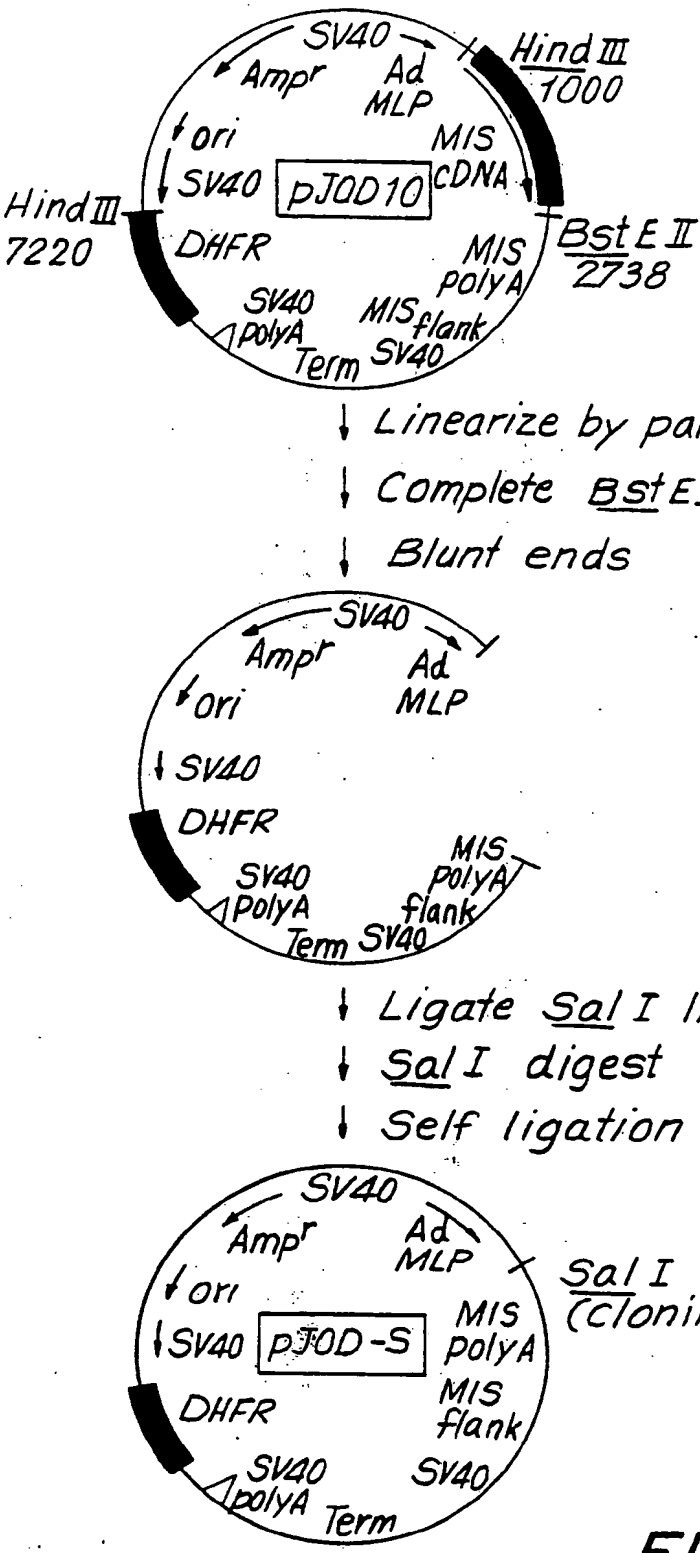

SV40
Ampr
Ad MLP
Hind III 1000
ori
SV40
pJOD10
MIS cDNA
Hind III 7220
DHFR
MIS polyA
Bst E II 2738
SV40 polyA Term
MIS flank SV40
Linearize by partial Hind III
Complete Bst E II
Blunt ends
SV40
Ampr
Ad MLP
ori
SV40
DHFR
MIS polyA flank
SV40 PolyA Term SV40
Ligate Sal I linkers
Sal I digest
Self ligation
SV40
Ampr
Ad MLP
Sal I (Cloning site)
ori
SV40
pJOD-S
MIS polyA
MIS flank
DHFR
SV40 polyA Term
SV40

FIG. 6E

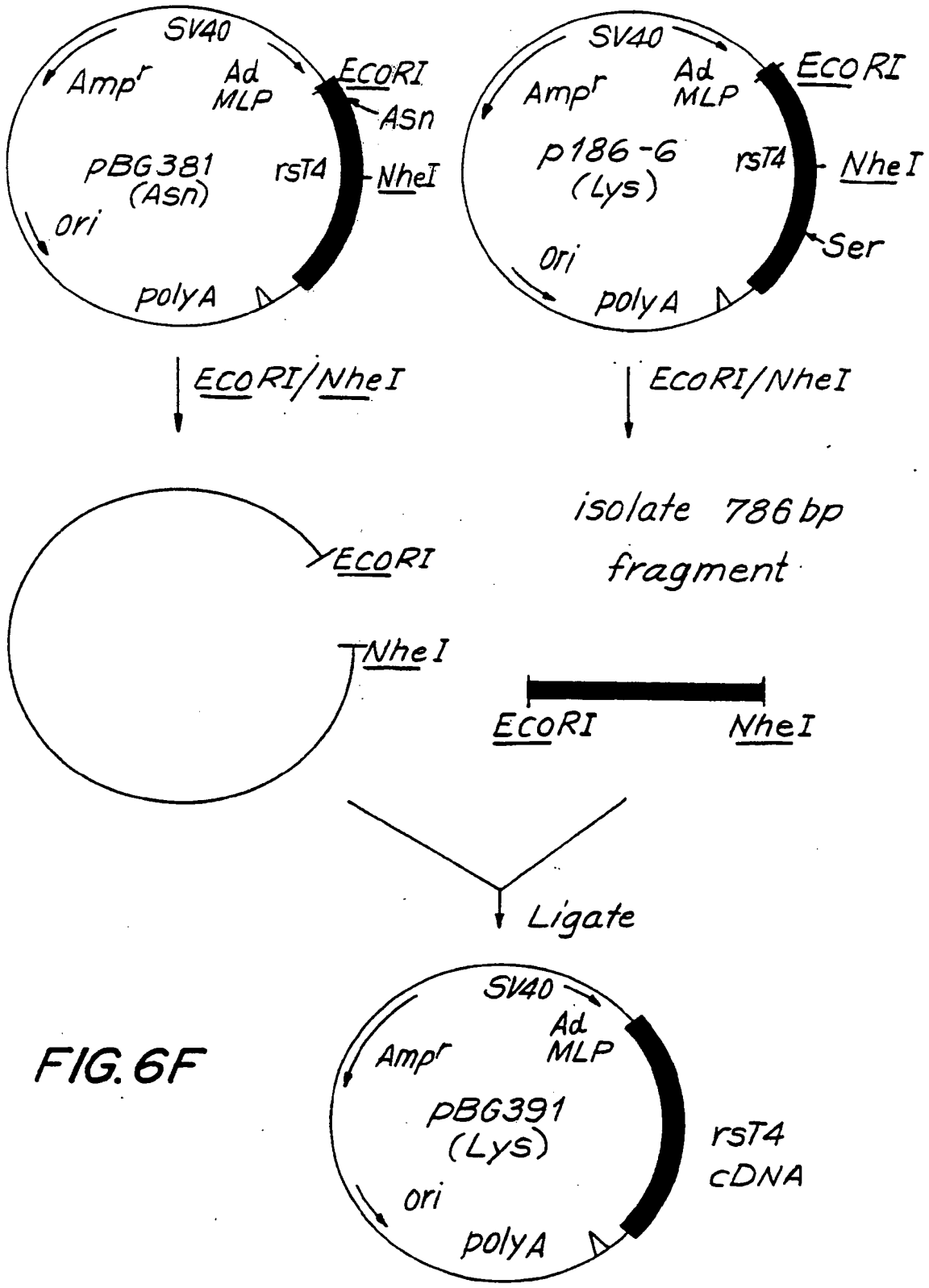

SV40
Amp^r
Ad MLP
EcoRI
Asn
pBG381 (Asn)
rsT4
NheI
ori
polyA
SV40
Amp^r
Ad MLP
EcoRI
p186-6 (Lys)
rsT4
NheI
Ser
ori
polyA
EcoRI/NheI
EcoRI/NheI
isolate 786 bp fragment
EcoRI
NheI
EcoRI
NheI
Ligate
SV40
Amp^r
Ad MLP
pBG391 (Lys)
rsT4 cDNA
ori
polyA

FIG. 6F

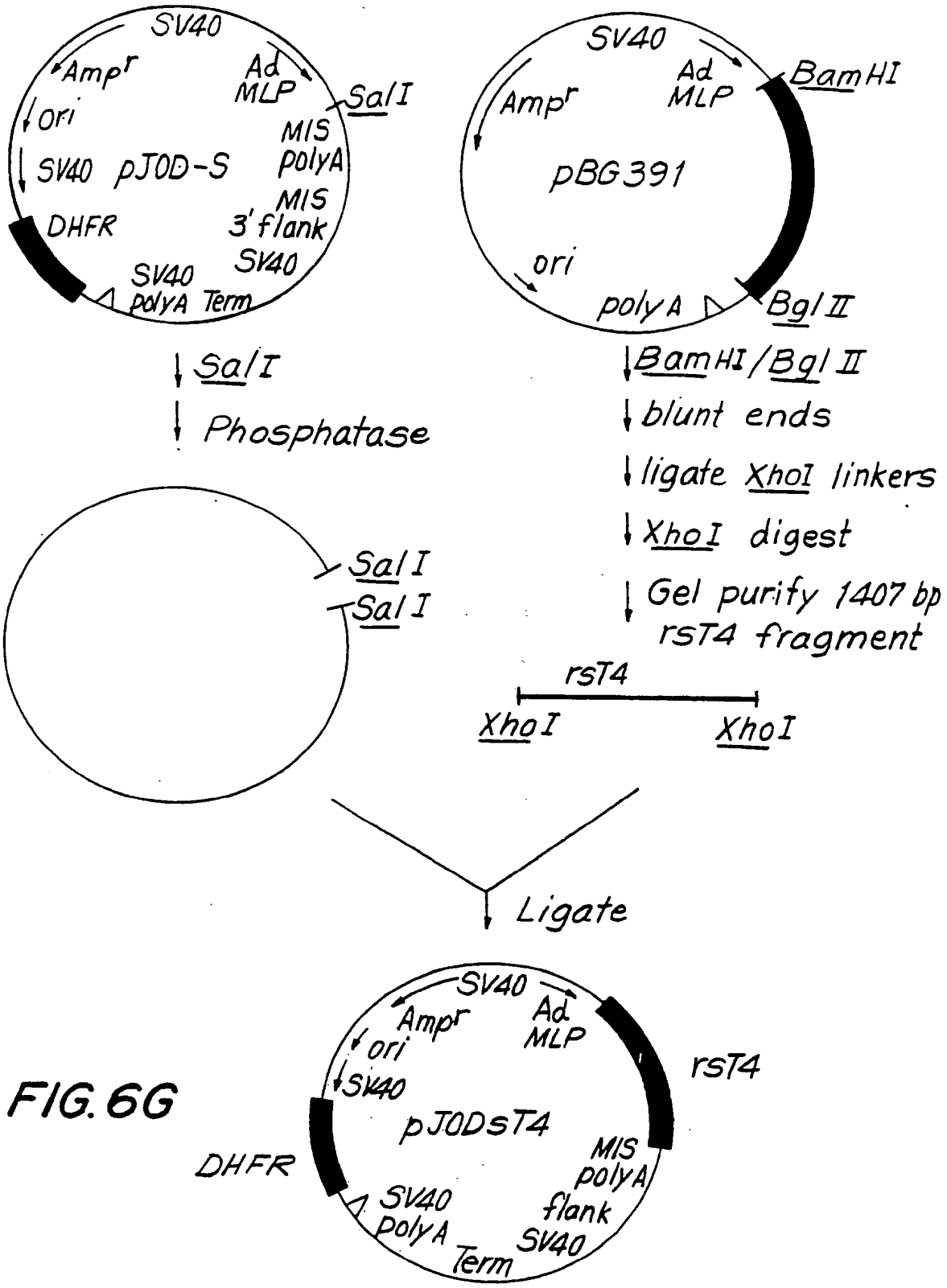

SV40
Ampr
Ad MLP
SalI
ori
MIS polyA
SV40 pJOD-S
MIS 3' flank SV40
DHFR
SV40 polyA Term
SV40
Ad MLP
BamHI
Ampr
pBG391
ori
polyA
BglII
SalI
Phosphatase
BamHI/BglII
blunt ends
ligate XhoI linkers
XhoI digest
Gel purify 1407 bp rsT4 fragment
SalI
SalI
rsT4
XhoI
XhoI
Ligate
SV40
Ampr
Ad MLP
ori
SV40
pJODsT4
rsT4
DHFR
MIS polyA flank SV40
SV40 polyA Term

FIG. 6G

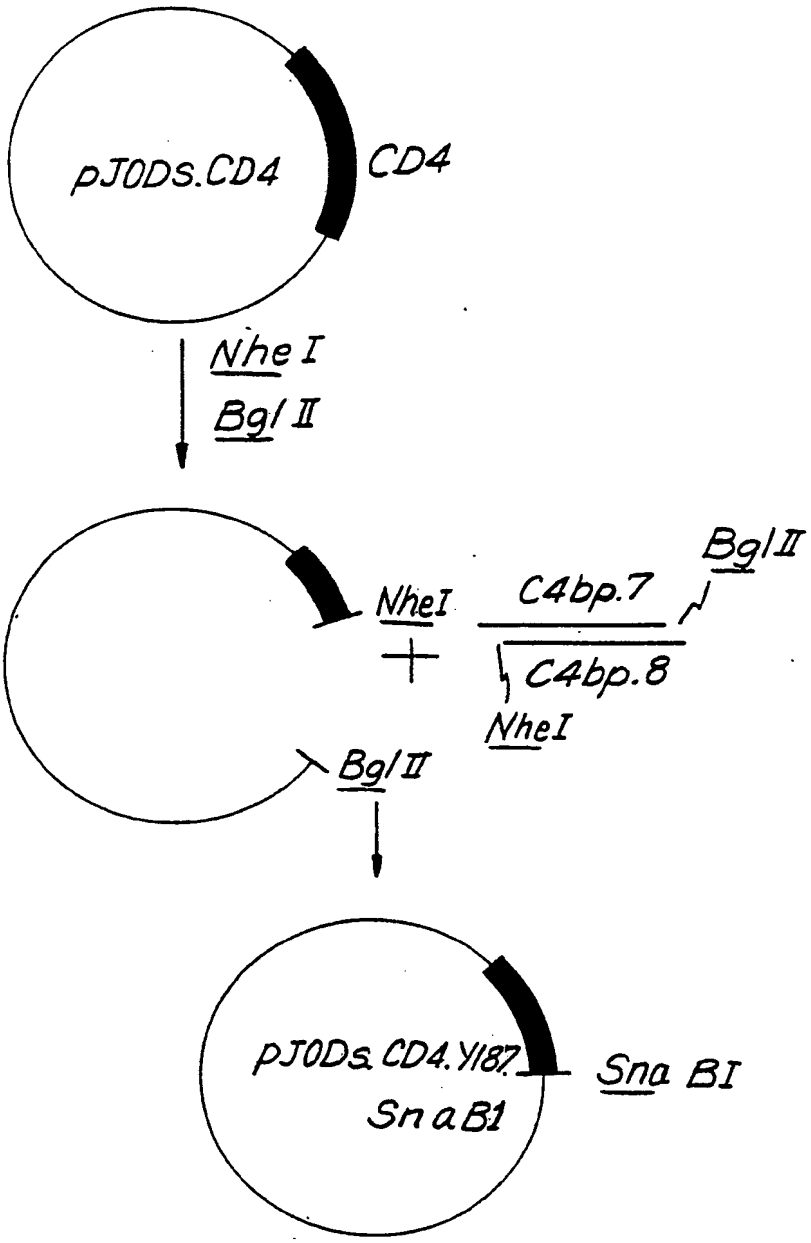
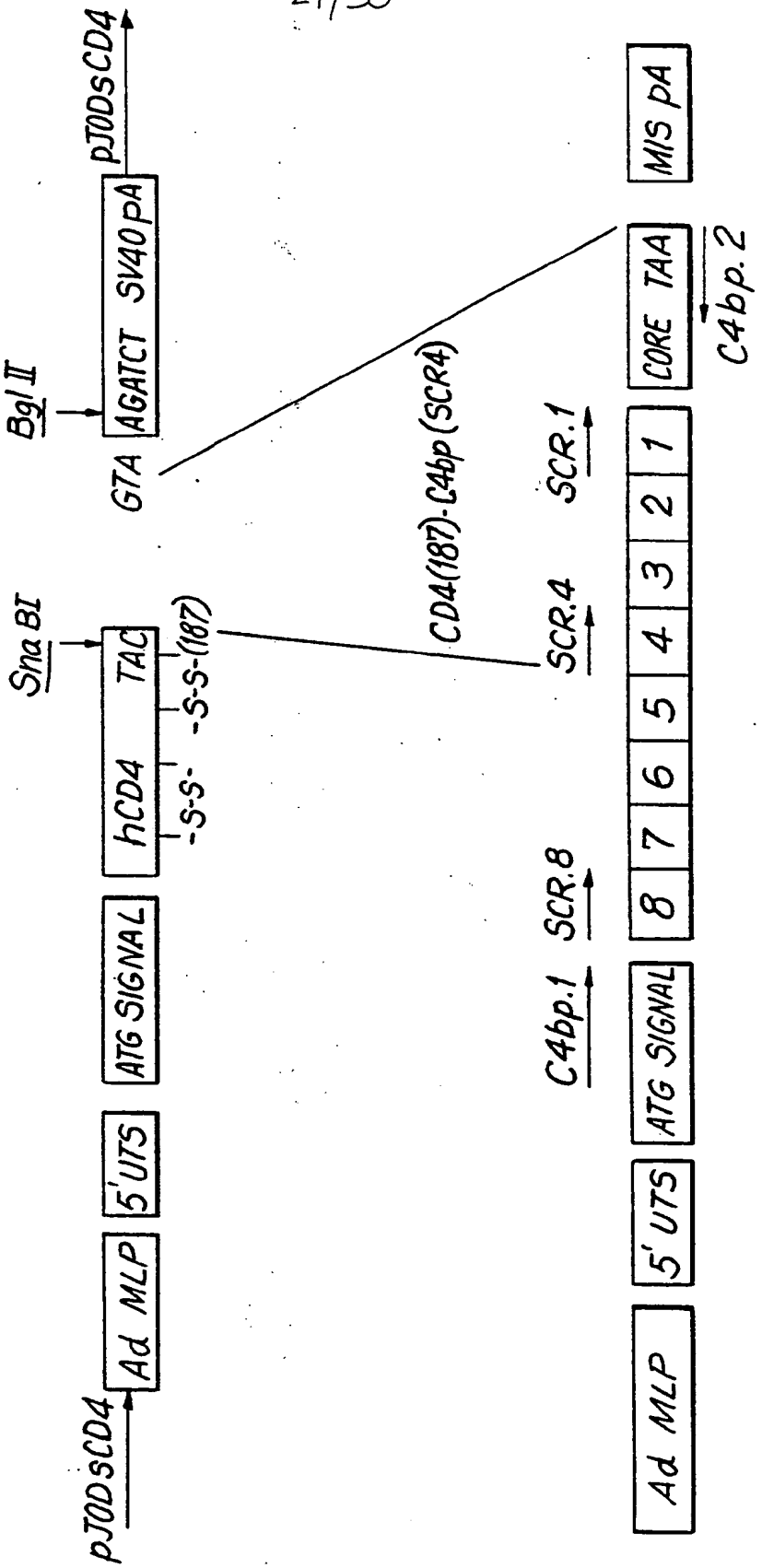

pJODs.CD4
CD4
Nhe I
Bgl II
NheI
C4bp.7
C4bp.8
BglII
NheI
BglII
pJODs. CD4.Y187.
Sna B1
Sna BI

FIG. 6H

FIG. 7



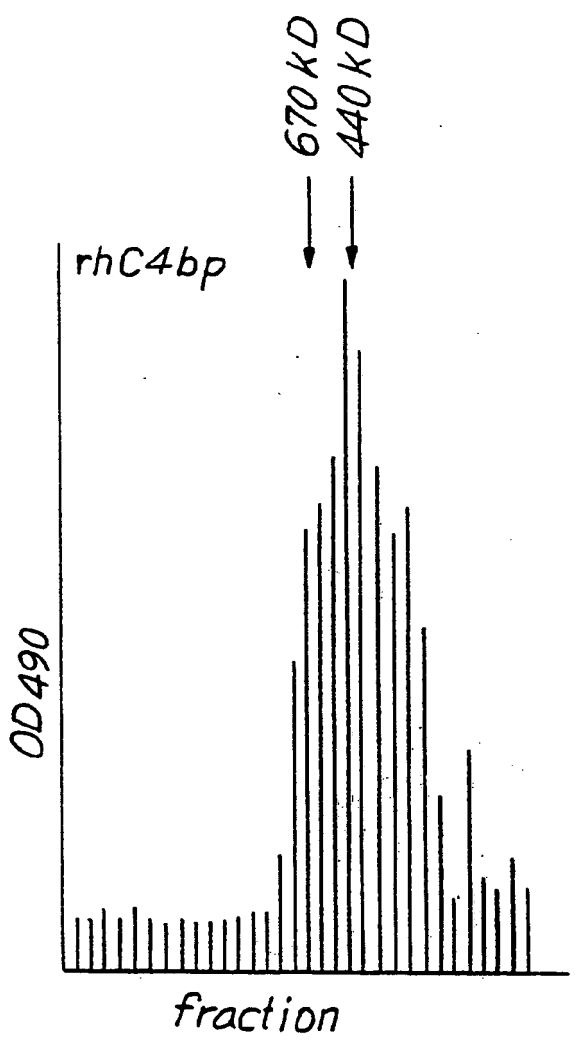


FIG. 8A

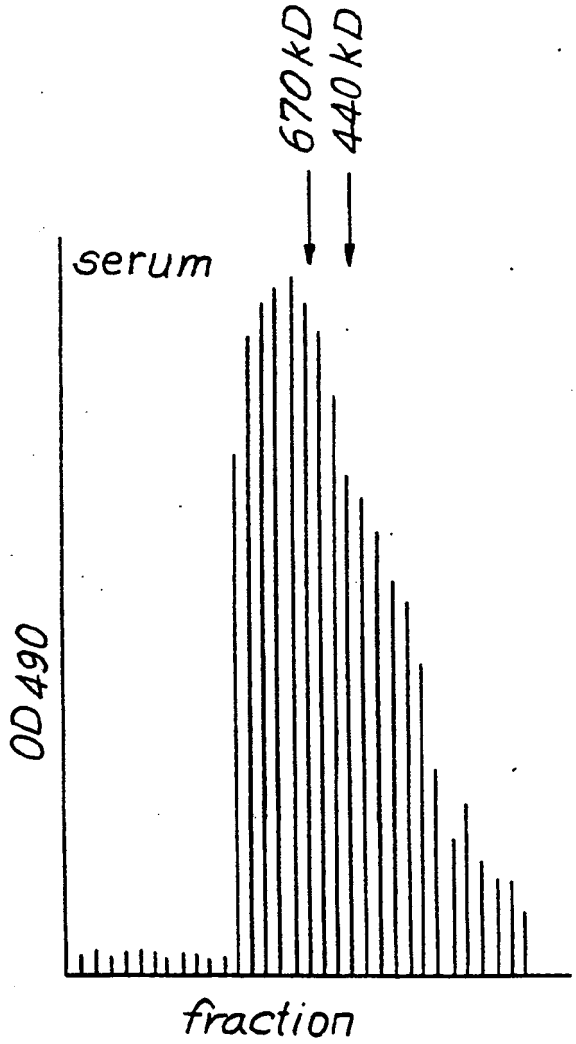


FIG. 8B

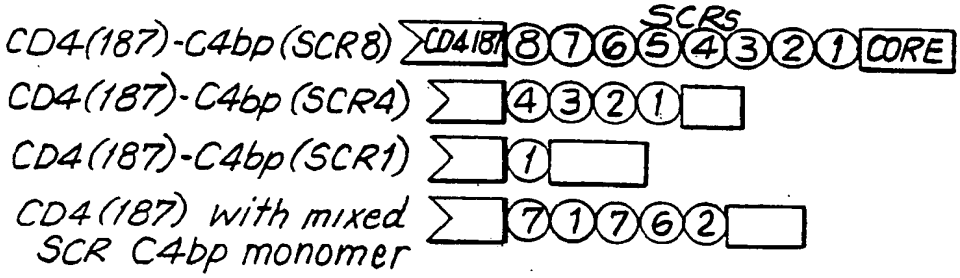


FIG. 9A

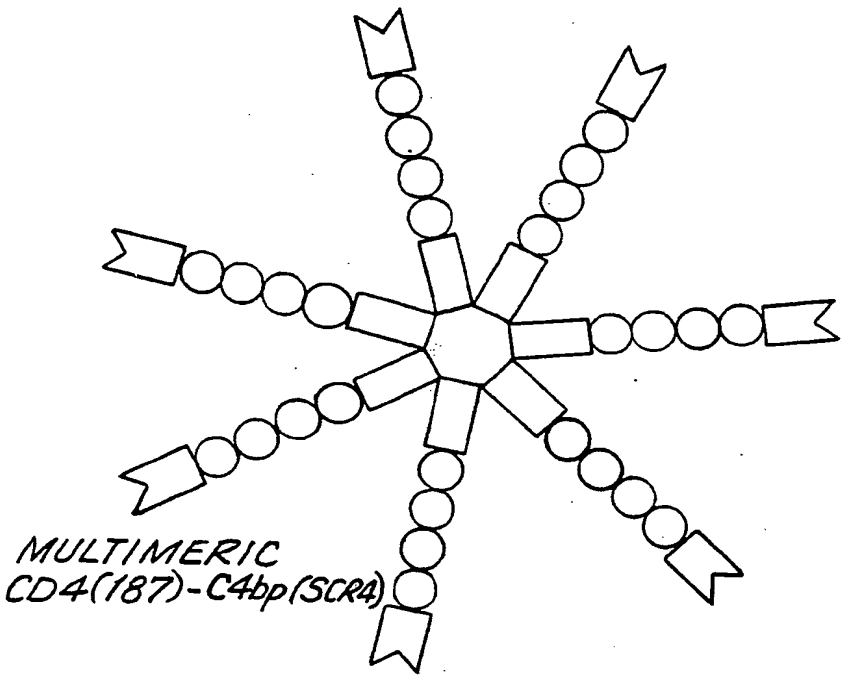


FIG. 9B

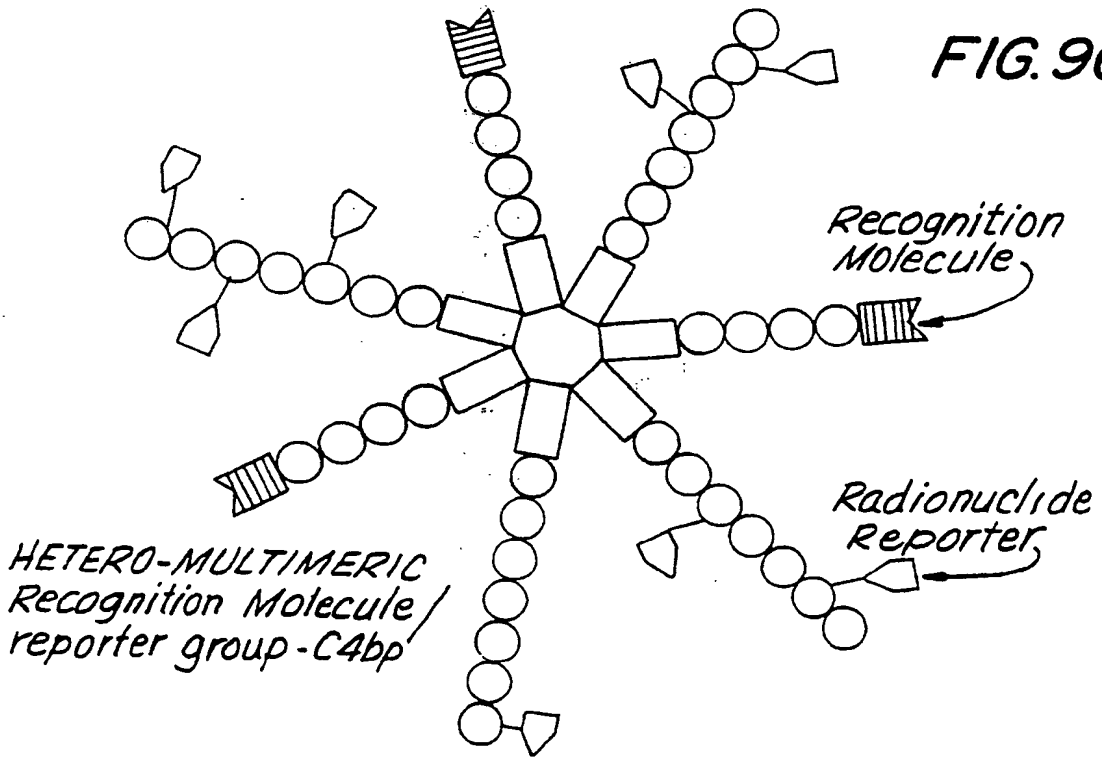


FIG. 9C

FIG. 10

ELISA #	Coating Antibody	Detection Antibody (HRP-conjugate)
ELISA 1	rabbit anti-hC4bp	6C6
2	6C6	6C6
3	gp120	rabbit anti-hC4bp/ goat anti-rabbit
4	gp120	6C6
5	gp120	5A8
6	gp120	051-198 or 051-28
7	1D7	5A8
8	rabbit anti-hC4bp	sheep anti-hC4bp
9	sheep anti-hC4bp	rabbit anti-hC4bp/ goat anti-rabbit

INTERNATIONAL SEARCH REPORT

International Application No PCT/US91/00567

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C07K 13/00; C12N 1/21, 7/01, 15/12; C12P 21/00
US: 530/409; 536/27; 435/320.1, 240.1, 69.7; 424/9; 514/12

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
US	536/27; 435/320.1, 240.1; 69.7; 530/409; 424/9; 514/12

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

Automated Patent Search, Chemical Abstracts Service database

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
A	Biochemistry, vol. 28, no. 11, issued 30 May 1989, Janatova et al. "Disulfide Bonds are localized within the Short Concensus Repeat Units of Complement Regulatory Proteins: C4b-Binding Proteins", pp. 4754-4761. See Abstract and Table I.	1-67
A	Proc. Natl. Acad. Sci, USA, vol. 84, issued December 1987, Bevilacqua et al. "Identification of an inducible endothelial-leukocyte adhesion molecule". pgs 9238-9242. See Introduction.	10,28,38
A	Nature, vol. 337, issued 09 February 1989, Capon et al. "Designing CD4 immunoadhesions for AIDS therapy", pg. 525-531. See entire article.	5-7,22-25, 35,36,42, 47-49,53

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
21 April 1991	29 MAY 1991
International Searching Authority [1]	**Signature of Authorized Officer [20]**
ISA/US	Nina Ossanna

Form PCT/ISA/210 (second sheet) (May 1986)